

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-249159

(43) 公開日 平成8年(1996)9月27日

(51) Int. Cl. <sup>6</sup>	識別記号	片内整理番号	F I	技術表示箇所
G 0 6 F 5/00			G 0 6 F 5/00	H A
12/08	3 2 0	7623-5B	12/08	3 2 0
13/10	3 4 0	7368-5E	13/10	3 4 0 A

審査請求 未請求 請求項の数 39 O L (全 46 頁)

(21) 出願番号 特願平7-49540

(22) 出願日 平成7年(1995)3月9日

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番  
1号

(72) 発明者 下井 洋行

神奈川県川崎市中原区上小田中1015番地  
富士通株式会社内

(72) 発明者 岡安 尚昭

神奈川県川崎市中原区上小田中1015番地  
富士通株式会社内

(74) 代理人 弁理士 竹内 進 (外1名)

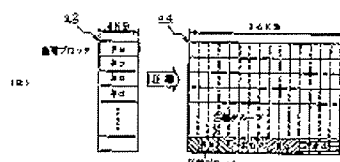
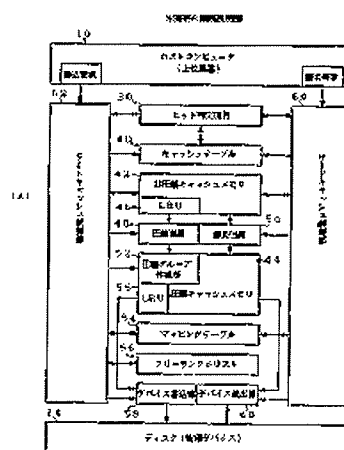
最終頁に続く

(54) 【発明の名称】 ディスク制御装置

(57) 【要約】

【目的】 ホストコンピュータ上のCPUの負荷を増加することなく、また、ディスク装置の実行的な使用率を低下させることなく、ディスク制御装置のハードウェアを用いてデータの圧縮と伸長及びデータの管理を行う。

【構成】 非圧縮キャッシュメモリ42から追い出された論理ブロックは圧縮回路48で圧縮され、圧縮グループ作成部52で圧縮データを複数集めた固定長の圧縮グループ(論理セクタサイズ)に纏められ、圧縮キャッシュメモリ44を経由してディスク14の論理セクタに格納される。マッピングテーブル54は、ディスク14の論理セクタに含まれる複数の論理ブロックとの対応関係を登録している。上位からの論理ブロック番号による要求に対し、マッピングテーブル54を参照してディスクの論理セクタ(圧縮グループ)を読み書きする。



(2)

特開平8-249159

1

2

## 【特許請求の範囲】

【請求項1】 所定サイズの論理セクタの割当てを受け、該論理セクタ単位にデータを読み書きする物理デバイスと、

上位装置と前記物理デバイスの間に設けられ、非圧縮データを上位装置の論理ブロック単位で格納する非圧縮キャッシュ領域と、圧縮データを前記物理デバイスの論理セクタと同一サイズをもつ圧縮グループ単位で格納する圧縮キャッシュ領域とを備えたキャッシュメモリと、上位装置の論理ブロック番号をインデックスとして、キャッシュ内アドレス、圧縮データが否かを示す圧縮フラグ、更新の有無を示す更新フラグを少なくとも登録し、前記キャッシュメモリの管理に使用されるキャッシュテーブルと、

上位装置からの要求ブロック番号のデータが前記キャッシュメモリに存在するか否かを判定するヒット判定部と、前記非圧縮キャッシュ領域に格納されたデータを前記論理ブロック単位に取り出して圧縮する圧縮回路部と、前記圧縮回路部による論理ブロック単位の圧縮データを集めて圧縮グループを作り、前記圧縮キャッシュ領域に格納する圧縮グループ作成部と、前記圧縮キャッシュ領域に格納されたデータを前記圧縮グループ単位に取り出して伸長し、前記非圧縮キャッシュ領域に展開する伸長回路部と、前記物理デバイスの空き論理セクタを管理するフリーリンクドリストと、

前記圧縮キャッシュ領域に格納されたデータを、前記圧縮グループ単位に取り出し、前記フリーリンクドリストから求めた空き状態にある論理セクタに書き込むデバイス書込部と、

前記上位装置の要求ブロック番号をインデックスとして、前記物理デバイスの格納位置を示す論理セクタ番号、論理セクタ内の相対位置、及びセクタ数を少なくとも登録し、前記物理デバイスの圧縮データの管理に使用されるマッピングテーブルと、

前記物理デバイスのデータを前記圧縮グループ単位に読み出して前記圧縮キャッシュ領域に転送するデバイス読出部と、

前記上位装置からの書込要求を処理するライトキャッシュ制御部と、

前記上位装置からの読出要求を処理するリードキャッシュ制御部と、を備えたことを特徴とするディスク制御装置。

【請求項2】 請求項1記載のディスク制御装置に於いて、前記圧縮回路部は、上位装置の論理ブロック単位に圧縮した圧縮データのサイズが、前記物理デバイスの最小物理セクタの整数倍でない場合は、前記最小物理セクタの整数倍となるようにダミーデータを付加することを特徴とするディスク制御装置。

【請求項3】 請求項1記載のディスク制御装置に於いて、

10

20

30

40

50

て、前記ライトキャッシュ制御部は、

前記上位装置からの書込要求に対し前記ヒット判定部でミスヒットを判定した場合、書込み論理ブロックを前記非圧縮キャッシュ領域に格納して上位装置に処理の終了を通知することを特徴とするディスク制御装置。

【請求項4】 請求項1記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、

前記上位装置からの書込要求に対し前記ヒット判定部で前記非圧縮キャッシュ領域でのヒットを判定した場合、前記非圧縮キャッシュ領域の該当する論理ブロックを上位装置からの書込み論理ブロックで更新し、上位装置に処理の終了を通知することを特徴とするディスク制御装置。

【請求項5】 請求項4記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、

前記非圧縮キャッシュ領域の論理ブロックを更新した際に、該更新した論理ブロックの属する圧縮グループの論理セクタ番号を前記マッピングテーブルから消去して前記フリーリンクドリストに移し、

更に、消去した前記論理グループに属する他の論理ブロックの圧縮データを前記グループ作成部に移して新たなグループ作成の組合せ候補とすることを特徴とするディスク制御装置。

【請求項6】 請求項1記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、

前記上位装置から論理ブロック番号を指定した書込要求に対し前記ヒット判定部で前記圧縮キャッシュ領域でのヒットを判定した場合、前記圧縮キャッシュ領域のヒットした圧縮データを前記伸長回路部で伸長して前記非圧縮キャッシュ領域に展開し、上位装置からの書込み論理ブロックで更新して上位装置に処理の終了を通知することを特徴とするディスク制御装置。

【請求項7】 請求項6記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、

前記非圧縮キャッシュ領域の論理ブロックを更新した際に、該更新した論理ブロックの属する圧縮グループの論理セクタ番号を前記マッピングテーブルから消去して前記フリーリンクドリストに移し、

更に、消去した前記論理グループに属する他の論理ブロックの圧縮データを前記グループ作成部に移して新たなグループ作成の組合せ候補とすることを特徴とするディスク制御装置。

【請求項8】 請求項1記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、

前記非圧縮キャッシュ領域からデータの追い出しを行った際に、前記キャッシュテーブルの参照で更新フラグのセットされた論理ブロックを前記追い出しデータから取り出し、前記圧縮回路で圧縮して前記グループ作成部に組合せ候補として供給することを特徴とするディスク制御装置。

(3)

特開平8-249159

3

4

【請求項9】請求項1記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、LRUに従って前記非圧縮キャッシュ領域からデータを追い出すことを特徴とするディスク制御装置。

【請求項10】請求項1記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、前記圧縮キャッシュ領域から圧縮グループの追い出しを行った際に、前記フリーリンクドリストの参照で空き状態にある論理セクタを割り当てて前記デバイス書込部により書き込ませ、

同時に前記マッピングテーブルに、書き込んだ圧縮グループに属する複数の論理ブロックの論理ブロック番号、論理セクタ番号、論理セクタ内の相対位置及びセクタ数の登録を行うことを特徴とするディスク制御装置。

【請求項11】請求項10記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、LRUに従って前記圧縮キャッシュ領域から圧縮グループを追い出すことを特徴とするディスク制御装置。

【請求項12】請求項1のディスク制御装置に於いて、前記圧縮グループ作成部は、前記圧縮データを前記圧縮回路部のバッファメモリに格納した順番で組み合わせて前記圧縮グループを作成することを特徴とするディスク制御装置。

【請求項13】請求項1のディスク制御装置に於いて、前記圧縮グループ作成部は、前記圧縮回路部で圧縮された順番に候補を組み合わせて前記圧縮グループを作成することを特徴とするディスク制御装置。

【請求項14】請求項1のディスク制御装置に於いて、前記圧縮グループ作成部は、前記圧縮回路部で圧縮した圧縮データのサイズを検知し、データサイズ毎に分類した候補リストを作成し、該候補リストの中から規定サイズを越える複数候補を選択して1つの圧縮グループを作成することを特徴とするディスク制御装置。

【請求項15】請求項14のディスク制御装置に於いて、前記圧縮グループ作成部は、前記サイズ毎に分類した候補リストの中から、同一サイズの候補を複数選択して1つの圧縮データグループを作成することを特徴とするディスク制御装置。

【請求項16】請求項14のディスク制御装置に於いて、前記圧縮グループ作成部は、前記サイズ毎に分類した候補リストの中から、異なるサイズの候補を複数選択して1つの圧縮データグループを作成することを特徴とするディスク制御装置。

【請求項17】請求項1のディスク制御装置に於いて、前記圧縮グループ作成部は、複数の候補を組み合わせた合計データサイズが規定の圧縮グループサイズに満たない場合は、残りの空き領域にダミーデータを格納することを特徴とするディスク制御

装置。

【請求項18】請求項1のディスク制御装置に於いて、前記圧縮グループ作成部は、1つの圧縮グループに組み合わせる候補の数を規定数に制限することを特徴とするディスク制御装置。

【請求項19】請求項1のディスク制御装置に於いて、前記圧縮グループ作成部は、1つの圧縮グループに組み合わせる候補の数を2個以下又は4個以下に制限することを特徴とするディスク制御装置。

10 【請求項20】請求項1のディスク制御装置に於いて、前記圧縮グループは、前記上位装置の論理ブロックの整数倍のサイズをもつことを特徴とするディスク制御装置。

【請求項21】請求項20のディスク制御装置に於いて、前記圧縮グループは、前記上位装置の論理ブロックの2<sup>n</sup>倍（但し、nは1、2、3、4、...の整数）のサイズをもつことを特徴とするディスク制御装置。

20 【請求項22】請求項1のディスク制御装置に於いて、前記デバイス書込部は、前記圧縮グループに、前記マッピングテーブルの情報を付加して前記物理デバイスに格納し、前記マッピングテーブルが故障した場合に前記マッピングテーブルを復元可能としたことを特徴とするディスク制御装置。

【請求項23】請求項1記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、前記圧縮グループの書込時に、ハッシングにより前記フリーリンクドリストから格納する前記物理デバイスの論理セクタを選択することを特徴とするディスク制御装置。

30 【請求項24】請求項23記載のディスク制御装置に於いて、前記ライトキャッシュ制御部は、前記圧縮グループの書込時に、同一シリンダアドレス内のハッシングにより前記フリーリンクドリストから格納する前記物理デバイスの論理セクタを選択することを特徴とするディスク制御装置。

40 【請求項25】請求項1記載のディスク制御装置に於いて、前記リードキャッシュ制御部は、前記上位装置からの読出要求に対し前記ヒット判定部でミスヒットを判定した場合、読出要求の論理ブロック番号による前記マッピングテーブルの参照で、前記物理デバイスから該当する論理セクタの圧縮グループを読み出して該当する圧縮データを取り出し、前記伸長回路で伸長して前記非圧縮キャッシュ領域に展開した後上位装置に転送して処理を終了することを特徴とするディスク制御装置。

【請求項26】請求項1記載のディスク制御装置に於いて、前記リードキャッシュ制御部は、前記上位装置からの読出要求に対し前記ヒット判定部で前記非圧縮キャッシュ領域でのヒットを判定した場合、前記非圧縮キャッシュ領域から該当する論理ブロックを上位装置に転送して処理を終了することを特徴とするデ

50

(4)

特開平8-249159

5

6

ディスク制御装置。

【請求項27】請求項1記載のディスク制御装置に於いて、前記リードキャッシュ制御部は、前記上位装置からの読出要求に対し前記ヒット判定部で前記圧縮キャッシュ領域でのヒットを判定した場合、前記圧縮キャッシュ領域から該当する論理ブロックを含む圧縮グループを取り出し、前記伸長部で伸長して前記非圧縮キャッシュ領域に展開して、該当する論理ブロックを上位装置に転送して処理を終了することを特徴とするディスク制御装置。

【請求項28】請求項1記載のディスク制御装置に於いて、前記リードキャッシュ制御部は、前記上位装置からの読出要求に対し前記ヒット判定部で前記圧縮キャッシュ領域でのヒットを判定した場合、前記圧縮キャッシュ領域から該当する論理ブロックを含む圧縮データのみ取り出し、前記伸長部で伸長して前記非圧縮キャッシュ領域に展開した後に上位装置に転送して処理を終了することを特徴とするディスク制御装置。

【請求項29】請求項1記載のディスク制御装置に於いて、前記キャッシュメモリ、キャッシュテーブル、ヒット判定部、圧縮部、圧縮グループ作成部、伸長部、デバイス書込部、マッピングテーブル、デバイス読出部、ライトキャッシュ制御部、及びリードキャッシュ制御部を1つのシステムとして二重化したシステムを構成し、前記上位装置は前記2つのシステムを共有し、前記物理デバイスは前記2つのシステムが共有することを特徴とするディスク制御装置。

【請求項30】請求項29記載のディスク制御装置に於いて、

前記二重化された2つのシステムの前記リードキャッシュ制御部の各々は、上位装置からの読出要求に対し独立に動作し、

前記二重化された2つのシステムの前記ライトキャッシュ制御部の各々は、自己のキャッシュメモリのミスヒット判定で上位装置からの新規データを格納した場合、該新規データを他のシステムのキャッシュメモリに複写すると共に他のシステムのキャッシュテーブルを更新し、また、自己のキャッシュメモリのヒット判定で上位装置からの新規データでヒットデータを更新した場合、該更新データを他のシステムのキャッシュメモリに複写すると共に他のシステムのキャッシュテーブルを更新することを特徴とするディスク制御装置。

【請求項31】請求項1記載のディスク制御装置に於いて、

前記圧縮グループ作成部は、更に、圧縮グループの作成時に各候補の先頭位置にグループ内での位置を示す位置フラグを付加し、該位置フラグは前記物理デバイスへの圧縮グループの書込み時に、前記マッピングテーブルに登録されることを特徴とするディスク制御装置。

【請求項32】請求項31記載のディスク制御装置に於

いて、

前記位置フラグは、グループ内の先頭位置、途中位置、グループ終端から次のグループ先頭までの途中位置のいずれかを表すことを特徴とするディスク制御装置。

【請求項33】請求項31記載のディスク制御装置に於いて、

前記リードキャッシュ制御部は、前記圧縮グループ内の位置フラグを使用して圧縮された論理ブロック単位の伸長を行うことを特徴とするディスク制御装置。

10 【請求項34】請求項1のディスク制御装置に於いて、前記キャッシュメモリは不揮発メモリであることを特徴とするディスク制御装置。

【請求項35】請求項1のディスク制御装置に於いて、前記マッピングテーブルを不揮発性メモリに格納したことを特徴とするディスク制御装置。

【請求項36】請求項1のディスク制御装置に於いて、前記物理デバイスが着脱自在な媒体を使用している場合、媒体排出に先立って前記マッピングテーブルを前記媒体に格納し、前記媒体投入時の初期化処理で前記媒体からマッピングテーブルを読み出して展開することを特徴とするディスク制御装置。

【請求項37】請求項1のディスク制御装置に於いて、前記物理デバイスは磁気ディスク装置、光ディスク装置又は半導体メモリ装置であることを特徴とするディスク制御装置。

【請求項38】請求項1のディスク制御装置に於いて、前記物理デバイスは複数のディスク装置を備えたディスクアレイ装置であることを特徴とするディスク制御装置。

30 【請求項39】所定サイズの論理セクタ単位にデータを読み書きする物理デバイスと、

圧縮データを前記物理デバイスの論理セクタと同一サイズをもつ圧縮グループ単位で格納する圧縮キャッシュ領域を備えたキャッシュメモリと、

上位装置の論理ブロック番号をインデックスとして、キャッシュ内アドレス、圧縮データか否かを示す圧縮フラグ、更新の有無を示す更新フラグが登録されるキャッシュテーブルと、

非圧縮データを前記論理ブロック単位に圧縮する圧縮回路部と、

40 前記圧縮回路部による論理ブロック単位の圧縮データを集めて前記圧縮キャッシュ領域に格納する圧縮グループ作成部と、

前記圧縮キャッシュ領域に格納されたデータを前記圧縮グループ単位に取り出して伸長する伸長回路部と、

前記物理デバイスの空き論理セクタを管理するフリーリンクドリストと、

前記圧縮キャッシュ領域に格納されたデータを、前記圧縮グループ単位に取り出し、前記フリーリンクドリストから求めた空き状態にある論理セクタに書き込むデバ

(5)

特開平8-249159

7

8

ス書込部と、

前記上位装置の要求ブロック番号をインデックスとして、前記物理デバイスの格納位置を示す論理セクタ番号、論理セクタ内の相対位置、及びセクタ数が登録されるマッピングテーブルと、

前記物理デバイスのデータを前記圧縮グループ単位に読み出して前記圧縮キャッシュ領域に転送するデバイス読出部と、

前記上位装置からの書込要求を処理するライトキャッシュ制御部と、

前記上位装置からの読出要求を処理するリードキャッシュ制御部と、を備えたことを特徴とするディスク制御装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、直接アクセス装置等の物理デバイスに対し圧縮データを記録再生するディスク制御装置に関し、特に、データサイズが変化する圧縮データであっても固定長ブロックデータとして物理デバイスとの間で記録再生を行うディスク制御装置に関する。

【0002】

【従来の技術】計算機システムの外部記憶装置として、マルチメディアに対応するために、多くのテキストファイル、プログラムファイル、画像データファイル等のデータを外部記憶装置内に格納する必要がある。外部記憶装置の特徴として、記録の不揮発性、大容量性、データ転送の高速性等の特徴をもつ磁気ディスク装置が広く用いられている。磁気ディスク装置の大容量、低価格を向上させる手法として、近年、データ圧縮が注目されている。

【0003】従来のデータ圧縮は、ソフトウェアを用いてホストコンピュータのソフトウェアでデータを圧縮することで、ディスク装置の容量を増大させていた。このソフトウェアによるデータ圧縮は、CPU性能を犠牲にしてディスク容量を増大させるものである。またソフトウェアによるデータ圧縮は、データのバックアップやデータの運搬等のように、データのアクセス応答速度をあまり気にしない分野に限られている。

【0004】データ圧縮をホストコンピュータ側で行う場合、ディスク装置に記憶されるデータ量は、ホストコンピュータで認識されるので、必要な記憶容量や残りの記憶容量は、圧縮処理がない場合と同様にホストコンピュータで管理できる。

【0005】

【発明が解決しようとする課題】しかし、ホストコンピュータの負荷を減らすために、ディスク制御装置のCPUでデータ圧縮を行うと、ディスク装置に記憶されるデータ量は、圧縮処理が終了するまで判明しないので、残りの記憶容量の管理がホストコンピュータで行えず、主記憶の確保に手間取る等の問題がある。

10

【0006】また、通常のホストコンピュータでは、物理ディスクの固定長セクタに対し、同じく固定長のデータブロック単位（論理ブロック単位）にデータの記録再生を行う。しかし、固定長の論理ブロックから得た圧縮データは、データの性格により種々のサイズの可変長データになる。そのため、現状の物理ディスクの固定長セクタに可変長の圧縮データを記録する方式では、圧縮データを有効に物理ディスク上に保管することはできず、物理ディスクの実効的な使用率が著しく低くなる。即ち、可変長の圧縮データを記録単位として物理ディスク上の固定長セクタに記録すると、データを記録しない無駄な部分が存在し、実効的なディスク使用率を低下させてしまうことが多い。

【0007】更にホストコンピュータからの書込み要求で物理ディスクから圧縮データを読み出して伸長した後更新した場合、更新後のデータを圧縮すると更新前に対しデータサイズが変化するため、必ずしも同じディスク場所に格納できない。この為、物理ディスク上にはフラグメンテーション（空き）が多数発生し、フラグメンテーションを無くするため、時々、セグメントクリーニングを実行しているときは、物理ディスクに対する入出力処理は停止しなければならない。その間、デバイスビジーとなり、アクセス性能が低下する問題がある。

30

【0008】更に、圧縮率が高くなると、上位装置で管理する論理アドレスと、圧縮データを管理する物理アドレスとの差が大きくなり、ディスク装置のシーク性能を考慮して決めていた論理アドレスの値が失われ、アクセス性能が低下する問題がある。従って、本発明の目的は、ホストコンピュータ上のCPUの負荷を増加することなく、また、ディスク装置の実効的な使用率を低下させることなく、ディスク制御装置のハードウェアを用いてデータの圧縮と伸長及びデータの管理を行うディスク制御装置を提供する。

【0009】

【課題を解決するための手段】図1は本発明の原理図である。ディスク装置等の物理デバイス14は、所定サイズの論理セクタ（最小物理セクタの整数倍）の割当てを受け、論理セクタ単位にデータを読み書きする。上位装置10と物理デバイス14の間には、キャッシュメモリが設けられる。キャッシュメモリは、非圧縮データを上位装置10の論理ブロック単位で格納する非圧縮キャッシュメモリ42と、圧縮データを物理デバイス14の論理セクタと同一サイズをもつ圧縮グループ単位で格納する圧縮キャッシュメモリ44に分けられる。

【0010】キャッシュテーブル30は、上位装置10の論理ブロック番号をインデックスとして、キャッシュ内アドレス、圧縮データか否かを示す圧縮フラグ、更新の有無を示す更新フラグを少なくとも登録し、キャッシュメモリの管理に使用される。ヒット判定回路30は、上

50

位装置10からの要求ブロック番号のデータがキャッシュメモリに存在するか否か判定する。

【0011】圧縮回路48は、非圧縮キャッシュメモリ42に格納されたデータを、論理ブロック単位に取り出して圧縮する。このとき得られる圧縮データのサイズはデータの性質により様々であり、必ずしも物理デバイス10、例えばディスクの初期化フォーマットで決まる最小アクセス単位である最小物理セクタ、例えば1KBの整数倍とはならない。

【0012】そこで、圧縮回路48は、上位装置10の論理ブロック単位に圧縮した圧縮データのサイズが、物理デバイス10の最小物理セクタの整数倍でない場合は、最小倍率セクタの整数倍となるようにダミーデータを付加する。例えば圧縮データが1.5KBであったならば、0.5KBのダミーデータを付加して2.0KBの圧縮ブロックとする。

【0013】圧縮グループ作成部52は、圧縮回路48による論理ブロック単位の圧縮データを集めて圧縮グループを作り、圧縮キャッシュメモリ44に格納する。伸長回路部50は、圧縮キャッシュメモリ44に格納されたデータを圧縮グループ単位に取り出して伸長し、非圧縮キャッシュメモリ42に展開する。フリーリンクドリスト56は、物理デバイス14の空き論理セクタを管理する。デバイス書込部58は、圧縮キャッシュメモリ44に格納されたデータを圧縮グループ単位に取り出し、フリーリンクドリスト56から求めた空き状態にある論理セクタに書き込む。マッピングテーブル54は、上位装置10の要求ブロック番号をインデックスとして、物理デバイス14の格納位置を示す論理セクタ番号、論理セクタ内の相対位置、及びセクタ数を少なくとも登録し、物理デバイス14の圧縮データの管理に使用される。

【0014】デバイス読出部60は、物理デバイス14のデータを圧縮グループ単位に読み出して圧縮キャッシュメモリ44に転送する。ライトキャッシュ制御部62は、上位装置10からの書込要求を処理する。リードキャッシュ制御部64は、上位装置10からの読出要求を処理する。ライトキャッシュ制御部62の処理は次のようになる。

【0015】まず上位装置10からの書込要求に対し、ヒット判定部30でミスヒットを判定した場合、書込み論理ブロックを非圧縮キャッシュメモリ42に格納して上位装置10に処理の終了を通知する。また上位装置10からの書込要求に対し、ヒット判定部30で非圧縮キャッシュメモリ42でのヒットを判定した場合、非圧縮キャッシュメモリ42の該当する論理ブロックを、上位装置10からの書込み論理ブロックに更新して上位装置10に処理の終了を通知する。このとき、更新した論理ブロックの属する圧縮グループの論理セクタ番号をマッピングテーブル54から消去してフリーリンクドリスト

に移す（圧縮グループの解散）。

【0016】更に、消去した論理グループに属する他の論理ブロックの圧縮データの圧縮データが圧縮キャッシュメモリ44に残っていたら、これを圧縮グループ作成部52に移し、新たなグループ作成の組合せ候補とする。更に、上位装置10からの書込要求に対し、ヒット判定部30で圧縮キャッシュメモリ44でのヒットを判定した場合、圧縮キャッシュメモリ44のヒットした論理ブロックを含む圧縮データを、伸長回路部で伸長して非圧縮キャッシュメモリ42に展開する。続いて非圧縮キャッシュメモリ42に展開した該当する論理ブロックを、上位装置10からの書込み論理ブロックに更新し、上位装置10に処理の終了を通知する。

【0017】この場合にも、更新した論理ブロックの属する圧縮グループの論理セクタ番号を、マッピングテーブル54から消去してフリーリンクドリスト60に移す。ライトキャッシュ制御部62は、LRU46に従って非圧縮キャッシュメモリ42からデータを追い出す。このとき追い出された論理ブロック番号でキャッシュテーブル40を参照し、更新フラグのセットされている論理ブロックを取り出し、圧縮回路48で圧縮して圧縮グループ作成部52に組合せ候補として供給する。

【0018】またライトキャッシュ制御部62は、LRU55に従って圧縮キャッシュメモリ44から圧縮グループを追い出す。追い出された圧縮グループに対し、フリーリンクドリスト56の参照で空き状態にある論理セクタを割り当て、デバイス書込部58により書き込ませる。同時にマッピングテーブル54に、書き込んだ圧縮グループに属する複数の論理ブロックの論理ブロック番号、論理セクタ番号、論理セクタ内の相対位置及びセクタ数を登録する。

【0019】圧縮グループ作成部52で作成する圧縮グループの候補の組合せは、圧縮データを圧縮回路のバッファメモリに格納されている順番、あるいは、圧縮された順番とする。また圧縮データのサイズを検知し、データサイズ毎に分類した候補リストを作成し、候補リストの中から規定サイズを超える複数候補を選択して1つの圧縮グループを作成してもよい。この場合、サイズ毎に分類した候補リストの中から、同一サイズの候補を複数選択して1つの圧縮データグループを作成してもよいし、異なるサイズの候補を複数選択して1つの圧縮データグループを作成してもよい。

【0020】圧縮データサイズによっては、複数の候補を組み合わせた合計データサイズが規定の圧縮グループサイズに満たない場合が起きる。この場合には、残りの空き領域にダミーデータを格納して1つの圧縮グループを作る。圧縮グループ作成部52で1つの圧縮グループに組み合わせる候補の数を規定数に制限することもできる。例えば1つの圧縮グループに組み合わせる候補の数を2個以下又は4個以下に制限する。このような制限は、

(7) 特開平 8-249159

11

ハードウェアで構成される圧縮回路 48 と伸長回路 50 の並列数に依存させる。例えば、圧縮回路 48 と伸長回路 50 を 2 回路設けている場合は、グループ内の圧縮データの数を 2 つに制限することで、並列処理を可能にして高速に圧縮または伸長できる。

【0021】本発明で使用する圧縮グループは、上位装置 10 の論理ブロックの整数倍のサイズをもつ。例えば平均的な圧縮率が  $1/2$  の場合は、2 倍のサイズが望ましい。また上位装置 10 の論理ブロックの  $2^n$  倍（但し、 $n$  は 1, 2, 3, 4, ... の整数）のサイズとしてもよい。デバイス書込部 58 は、更に、論理セクタの先頭、マッピングテーブル 54 に登録する情報をセクタ情報として付加して物理デバイス 14 に格納する。これにより、マッピングテーブル 54 が故障した場合に、論理セクタのセクタ情報を物理デバイス 14 から読み出してマッピングテーブル 54 を再構成できる。

【0022】ライトキャッシュ制御部 62 は、圧縮グループの書込時に、ハッシングによりフリーリンクドリスト 54 から格納する物理デバイス 54 の論理セクタを選択してもよい。例えば同一シリンダアドレス内のハッシングによりフリーリンクドリスト 56 から格納する物理デバイス 14 の論理セクタを選択する。これによってディスク装置のシーク性能を高める。

【0023】リードキャッシュ制御部 64 の処理は次のようになる。まずリードキャッシュ制御部 64 は、上位装置 10 からの読出要求に対し、ヒット判定部 30 でミスヒットを判定した場合、読出要求の論理ブロック番号によるマッピングテーブル 54 の参照で、物理デバイス 14 から該当する論理セクタの圧縮グループを読み出し、その中から該当する圧縮データを取り出して伸長回路 50 で伸長してキャッシュメモリ 42 に展開し、上位装置 10 に転送して処理を終了する。

【0024】また上位装置 10 からの読出要求に対し、ヒット判定部 30 で非圧縮キャッシュメモリ 42 でのヒットを判定した場合、非圧縮キャッシュメモリ 44 から該当する論理ブロックを上位装置 10 に転送して処理を終了する。更に、上位装置からの読出要求に対し、ヒット判定部 30 で圧縮キャッシュメモリ 44 でのヒットを判定した場合、圧縮キャッシュメモリ 44 から該当する論理ブロックを含む圧縮グループを取り出し、伸長回路 50 で伸長して非圧縮キャッシュメモリ 42 に展開し、該当する論理ブロックを上位装置 10 に転送して処理を終了する。

【0025】さらに本発明は、キャッシュメモリ、キャッシュテーブル 40、ヒット判定部、圧縮回路 48、圧縮グループ作成部 52、伸長回路 50、デバイス書込部 58、マッピングテーブル 54、デバイス読出部 60、ライトキャッシュ制御部 62、及びリードキャッシュ制御部 62 を 1 つのシステムとして二重化したシステムを構成し、上位装置は 2 つのシステムを共有し、物理デバ

12

イス 14 は 2 つのシステムが共有する。

【0026】二重化された 2 つのシステムのリードキャッシュ制御部 64 の各々は、上位装置 10 からの読出要求に対し独立に動作する。また二重化された 2 つのシステムのライトキャッシュ制御部 62 の各々は、自己のキャッシュメモリのミスヒット判定で上位装置 10 からの新規データを格納した場合、新規データを他のシステムのキャッシュメモリに複写すると共に他のシステムのキャッシュテーブルを更新する。また、自己のキャッシュメモリのヒット判定により上位装置からの新規データでヒットデータを更新した場合にも、更新データを他のシステムのキャッシュメモリに複写すると共に他のシステムのキャッシュテーブル 40 を更新する。

【0027】更に本発明の他の変形として、圧縮グループ作成部 52 は、圧縮グループの作成時に各候補の先頭位置にグループ内での位置を示す位置フラグを付加し、この位置フラグは物理デバイス 14 への圧縮グループの書込み時に、マッピングテーブル 54 にも登録する。位置フラグは、グループ内の先頭位置、途中位置、グループ終端から次のグループ先頭までの途中位置のいずれかを表す。リードキャッシュ制御部 64 は、圧縮グループ内の位置フラグを使用して圧縮された論理ブロック単位の伸長を行う。

【0028】キャッシュメモリには不揮発メモリを使用する。マッピングテーブル 54 は、キャッシュメモリとしても設けた不揮発性メモリに格納する。物理デバイス 14 は、磁気ディスク装置、光ディスク装置又は半導体メモリ装置である。また物理デバイス 14 は、複数のディスク装置を備えたディスクアレイ装置でもよい。

【0029】

【作用】このような本発明の圧縮データ記憶システムは次の作用を有する。まずデータを圧縮した後に問題となるのがデータ長である。つまり、ホストコンピュータが管理するデータは、通常、固定長のデータブロックであるが、固定長データブロックを圧縮すると、データ長は変化してしまう。これは、データの性格に依存して、圧縮データのサイズが決まるためである。また、ホストコンピュータから磁気ディスク装置を見た場合、従来と同様に、非圧縮のデータブロックを認識したデータ管理が必要である。これはデータの等価性を保持するためである。その為に、ホストコンピュータが管理する論理ブロック番号（論理ブロックアドレス）と、磁気ディスク装置が管理する論理セクタ番号（論理アドレス番号）の変換用テーブルであるマッピングテーブルがディスク制御装置には必要になる。

【0030】そこで、本発明は、ホストコンピュータがデータの記憶場所を管理するのではなく、ディスク制御装置が可変長の圧縮データを、固定長の圧縮グループ（論理グループ）に纏めた形で管理する方法を採用する。即ち、一定サイズの論理セクタに、固定長ブロック

(8)

特開平 8-249159

13

14

データを圧縮して得た圧縮データを複数収めることにより、可変長の圧縮データを固定長データとして扱い、物理ディスク上の空白の部分が多く残らないようにデータを記録し、必要なデータ管理を行う。

【0031】非圧縮のデータは、上位装置の論理ブロックサイズで管理され、論理ブロック単位に圧縮される。このように論理ブロックから圧縮された圧縮データを多数集め、ある一定の論理セクタサイズをもつサイズの論理グループになるように選択してから、物理ディスクの論理セクタを付与してデータを記録する。ここでディスク装置に記録できる最小単位は、物理ディスクを初期化した際の物理セクタサイズで決まる。論理セクタは最小物理セクタの整数倍のサイズとする。選択した圧縮データが論理セクタのサイズに満たない場合は、ダミーデータを付加する1論理セクタ分の圧縮データとする。

【0032】ディスク装置に対する論理セクタの格納に際しては、論理セクタ番号、論理グループ内の圧縮データの論理ブロック番号、圧縮フラグ、先頭論理ブロック番号からの相対位置、セクタ数などの対応関係をマッピングテーブルに登録する。このマッピングテーブルに登録する情報を、論理セクタの先頭にセクタ情報として付加してディスク装置に格納してもよい。

【0033】ディスク制御装置は、上位装置から受信した論理ブロック番号によりマッピングテーブルを参照して論理セクタの位置を検出する。そのためディスク制御装置では、マッピングテーブルは最も重要なものと言える。マッピングテーブルが故障した場合には、論理セクタの先頭に付加しているセクタ情報を読み出すことによって、マッピングテーブルを再構成することができる。

【0034】上位装置の書き要求で特定の論理ブロックの更新が行われた場合、更新データが含まれていた論理グループは解体され、キャッシュメモリ上で新規な論理グループを形成する候補となり、必ず新たな論理グループに組み合わされて格納場所も変化する。またグループ解体時にはマッピングテーブルの論理セクタの登録も消去して開放する。このためディスク装置のセグメントクリーニングの必要は無く、セグメントクリーニングのためにデバイスビジーとなることはない。

【0035】

【実施例】

<目次>

1. 動作環境
2. 機能ブロックの構成
3. データの記録処理
4. データの読出処理
5. 二重化構成
6. 二重化システムの書き込み処理
7. 二重化構成の読出処理
8. 位置フラグの付加による圧縮データのブロック化
9. 他の動作環境

## 1. 動作環境

図2は、本発明の圧縮データ記憶システムが実施される動作環境のブロック図である。図2において、ホストコンピュータ10に対し、磁気ディスク制御装置12と複数のディスク装置14-1～14-6を備えたディスクアレイによって、外部記憶のためのサブシステムを構成している。

【0036】磁気ディスク制御装置12には、MPU16、ROM18、RAM20、ホストインタフェース制御装置22、ディスクインタフェース制御装置24、キャッシュ制御部26、キャッシュメモリ28、ヒット判定回路30および圧縮伸長回路32が設けられ、それぞれ内部バス34によりMPU16に接続されている。MPU16は、ROM18に内蔵されたプログラムにより磁気ディスク制御装置12全体を制御する。特に本発明にあつては、磁気ディスク制御装置12に圧縮伸長回路32を備えており、ホストコンピュータ36からのデータを圧縮伸長回路32で圧縮してディスク装置14-1～14-6に格納し、またディスク装置14-1～14-6から読み出した圧縮データを伸長してホストコンピュータ10に転送するデータ圧縮と復元の制御を行う。

【0037】更にMPU16は、ホストコンピュータ10からの論理ブロック単位のデータを圧縮した後に、所定の論理セクタにまとめた圧縮グループを作成してディスク装置14-1～14-6に書き込むための制御も行う。ここでホストインタフェース制御部22およびディスクインタフェース制御装置24は、例えばSCSIを想定している。

【0038】キャッシュメモリ28は、ホストコンピュータ10と物理デバイスとしてのディスク14-1～14-6との間で転送されるデータを一時的に保持する。またキャッシュメモリ28は、圧縮グループにまとめられた圧縮データも一時的に保持する。圧縮伸長回路32でホストコンピュータ10の論理ブロック単位に圧縮された圧縮データは、キャッシュメモリ28内で一定サイズの圧縮グループにまとめられた後、物理デバイスとしてのディスク装置14-1～14-6に格納される。

【0039】圧縮伸長回路32は、後の説明で明らかにするように、バッファメモリを内蔵しており、バッファメモリに転送されたデータの圧縮と伸長を行う。この圧縮伸長回路32に内蔵されたバッファメモリとキャッシュメモリ28との間のデータ伝送は、MPU16が制御する。磁気ディスク制御装置12は、物理的に接続されている複数のディスク装置14-1～14-6を1つの論理ディスクとして認識している。この論理ディスクには、圧縮データがまとめられて圧縮グループ単位に格納される。ここでディスク装置14-1～14-6の合計ディスク容量を1GBとすると、磁気ディスク制御装置12が認識する論理ディスクの容量は1GBとなるが、データの圧縮によりディスク容量が2倍になることを予

15

定すると、論理ディスクの容量は2GBに拡張して扱う。

【0040】更に、ディスク装置14-1~14-6をアクセス可能な最小セクタ単位は、初期化処理のフォーマットで決まっており、例えば最小物理セクタは1KBとなる。そこで以下の説明にあたっては、ホストコンピュータ10が扱っている論理ブロックのサイズを4KBとし、圧縮伸長回路32はこの4KBの論理ブロックサイズを1単位として圧縮および伸長を行うものとする。また論理ブロック単位の圧縮で得られた複数の圧縮データをまとめた圧縮グループ（圧縮クラスタともいう）のサイズを、論理ブロックの4倍の16KBとする。

【0041】図3は、図2の動作環境の具体例として、RAIDコントローラへの適用を示す。RAIDコントローラは、ホストIF用プロセッサモジュール100、ホストIFモジュール102、デバイスIF用プロセッサモジュール106、デバイスIFモジュール108-1~108-6で構成される。このような通常のRAIDコントローラに対し、本発明にあたっては、更にデータ圧縮伸長モジュール104を追加している。

【0042】ホストIF用プロセッサモジュール100は、メインプロセッサ110、コ-プロセッサ112、フラッシュメモリ114、プロセッササポートロジック116およびシステムストレージ118で構成される。この構成は、デバイスIF用プロセッサモジュール106についても同じであり、メインプロセッサ136、コ-プロセッサ138、フラッシュメモリ140、プロセッササポートロジック142およびシステムストレージ144で構成される。

【0043】ホストIF用プロセッサモジュール100とデバイスIF用プロセッサモジュール106は、プロセッサバスブリッジ156を介してコントロールバスにより接続されている。ホストIFモジュール102は、ホストインタフェース120、ホストバッファインタフェース122、キャッシュメモリ124を有する。このキャッシュメモリ124に対してはバックアップ電源125が設けられており、不揮発性のキャッシュメモリを構成している。

【0044】デバイスIFモジュール108-1~108-6は、デバイスIFモジュール108-1に代表して示すように、デバイスバッファIF146、バッファメモリ148およびデバイスIF150を有する。このようなRAIDコントローラに新たに追加したデータ圧縮伸長モジュール104は、上位インタフェース126、下位インタフェース128、伸長データ格納用バッファメモリ130、圧縮データ格納用バッファメモリ132および4つの伸長圧縮ロジック134-1~134-4を有する。

【0045】図4は、図3のデータ圧縮伸長モジュール104の詳細である。上位インタフェース126は、ホ

(9)

特開平8-249159

16

スト側の内部バスとの接続に設けられる。下位インタフェース128は、デバイス側のインタフェースとの接続に用いられる。上位インタフェース126には、圧縮LSI制御回路160、システムバス制御回路162、制御レジスタ164、アドレスレジスタ166、データレジスタ168、アービタ170およびコマンドレジスタ172が設けられている。この内部構成は、下位インタフェース128についても同じである。

【0046】圧縮伸長ロジック134-1~134-4は、圧縮伸長ロジック134-1に代表して示すように、データ圧縮回路172、データ伸長回路174、辞書176、比較部178を備える。圧縮伸長ロジック134-1~134-4で実現される圧縮アルゴリズムは、算術符号化アルゴリズムA2（アルゴリズム）、L2Wアルゴリズムなど適宜のアルゴリズムを適用することができる。

【0047】特に図4の実施例にあたっては、4回路の圧縮伸長ロジック134-1~134-4を備えていることから、最大で4並列の圧縮処理または伸長処理を行うことができる。本発明にあたっては、伸長データ格納バッファメモリ130にキャッシュメモリから追い出しを受けたデータを格納し、圧縮データ格納用バッファメモリ132に格納する。圧縮データ格納用バッファメモリ132にあたっては、複数の圧縮データをまとめて例えば16KBの圧縮グループ（圧縮クラスタ）を作成している。複数の圧縮データを用いた圧縮グループの作成は、後の説明で明らかにする。圧縮データ格納用バッファメモリ132で複数の圧縮データを組み合わせて圧縮グループが作成されると、この圧縮グループはキャッシュメモリに展され、その後に論理ディスクを構成するアレイ構成のディスク装置に格納される。

【0048】再び図3を参照するに、このRAIDコントローラにあたっては、図示のモジュール側をシステム1とすると、同様なモジュール構成によるシステム2を設けることで二重化構成としている。システム1とシステム2の結合は、ブリッジバッファインタフェース152と周辺ロジック154を介したデータバスの結合で行える。このため、システム1とシステム2の間でデータおよび各種のコマンドのやり取りが可能となる。

## 2. 機能ブロックの構成

図5は、図2の磁気ディスク制御装置12の機能ブロックである。本発明の磁気ディスク制御装置12を機能的に見ると、ヒット判定回路30、キャッシュテーブル40、非圧縮キャッシュメモリ42、圧縮キャッシュメモリ44、圧縮回路48、伸長回路50、圧縮グループ作成部52、マッピングテーブル54、フリーリンクドリスト56、デバイス書込部58、デバイス読出部60、ライトキャッシュ制御部62およびリードキャッシュ制御部64を備える。

【0049】更に、キャッシュ機能を実現するため非圧

(10)

特開平8-249159

17

縮キャッシュメモリ42にLRU部46が設けられ、同様に圧縮キャッシュメモリ44にLRU部55が設けられている。図6は図2のキャッシュメモリ28であり、キャッシュメモリ28は非圧縮データ領域42と圧縮データ領域44に分けられている。図5の機能ブロックにあっては、非圧縮データ領域42を非圧縮キャッシュメモリ42としており、また圧縮データ領域44を圧縮キャッシュメモリ44としている。このキャッシュメモリ28が、図3のハードウェア構成に示したように、バックアップ電源125を有することから不揮発メモリとなっている。更に、不揮発メモリとしたキャッシュメモリ28にはマッピングテーブル54を格納している。

【0050】キャッシュメモリ28の非圧縮データ領域42にあっては、ホストコンピュータ10の処理単位である4KBの論理ブロックサイズをキャッシュブロック（キャッシュセグメント）としてデータを管理する。一方、圧縮データ領域44については、16KBの圧縮グループ単位にキャッシュブロック（キャッシュセグメント）を管理している。

【0051】キャッシュメモリ28の内容は、キャッシュテーブル40に格納されている。図7は、本発明のキャッシュテーブル40の説明図である。キャッシュテーブル40は、ホストコンピュータの扱う論理ブロック番号40-1ごとにテーブル情報が作成され、論理ブロック番号40-1に対応して格納位置を示すキャッシュ内のアドレス40-2、データ長40-3、圧縮データか否かを示す圧縮フラグ40-4および、キャッシュ上での更新（書替え）が行われたか否かを示す更新フラグ40-5を登録している。

【0052】このため、ヒット判定回路30はホストコンピュータ10からの論理ブロック番号を指定した書込要求または読出要求に対し、図7のキャッシュテーブル40を参照して、登録された論理ブロック番号40-1の有無でキャッシュのミスヒットとヒットを判定することができる。また、ヒットした場合には、該当する論理ブロック番号40-1に対応するテーブルデータからキャッシュ内のアドレス40-2、データ長40-3、圧縮フラグ40-4がセットの際は圧縮データ、リセットの場合は非圧縮データ、更新フラグ40-5がセットの場合は更新済み、リセットの場合は未更新を知ることができる。

【0053】図8は、図5の圧縮回路48によるキャッシュメモリでの非圧縮データと圧縮データのやり取りである。まず非圧縮キャッシュメモリ42は、ホストコンピュータ10の論理ブロックサイズに一致する4KBのサイズをもつキャッシュブロック68-1～68-nでデータを管理している。今、非圧縮キャッシュメモリ42には、論理ブロック番号#a、#b、#c、#dのデータが格納されていたとする。この論理ブロック番号#a、#b、#c、#dがLRUにより非圧縮キャッシュ

18

メモリ42から追い出されると、圧縮回路48はキャッシュブロック68-1～68-4のそれぞれを圧縮し、圧縮データを作り出す。

【0054】ここで圧縮キャッシュメモリ44は、4KBのキャッシュブロック68-1～68-nに対し4倍の16KBの大きさをもつ圧縮グループ（圧縮クラス）単位で管理されている。このため、論理ブロック番号#a～#dのキャッシュブロック68-1～68-4の圧縮データ72-1～72-4を、例えば圧縮グループ70-nのように1つのデータブロックにまとめる。

【0055】図9は、圧縮回路48による4KBの論理ブロックの圧縮データである。今、図9（A）に示す4KBの論理ブロック68を圧縮回路48で圧縮し、図9（B）の圧縮データ72を得たとする。この圧縮データは例えば1.5KBバイトであったとする。本発明の物理デバイスとなるディスク装置14-1～14-6は、初期化時のフォーマットによりアクセス可能な最小物理セクタの大きさが例えば1KBと決まっている。

【0056】このため、1.5KBの圧縮データ72のままでは、最小物理セクタが1KBのディスク装置に格納することができない。そこで、圧縮データ72の後ろに0.5KBのダミーデータ82を付加して2KBの圧縮ブロックを作り出す。このため、圧縮ブロックは最低でもディスク装置の最小物理セクタに等しい1KBのサイズをもつことになる。

【0057】このように、必要に応じてダミーデータ82を付加していた圧縮ブロックは、下側に取り出して示すように、スタートコード84に続いて、本来の圧縮データと必要に応じて設けたダミーデータを含む圧縮データ部86を有し、末尾にエンドコード88を設けている。このような圧縮ブロックを作ることによって、圧縮回路48および伸長回路50は、最小物理セクタの整数倍のサイズをもつ論理ブロック番号に対応した圧縮ブロックの組合せによるグループ化と、グループの中からの必要な圧縮ブロックの取出しが可能となる。

【0058】次に図5の圧縮グループ作成部52の処理を説明する。図10は、論理ブロック単位の圧縮で得られた圧縮ブロックから圧縮グループを作成する処理の一例である。この圧縮グループの作成にあっては、圧縮回路48で論理ブロックを圧縮する際に、作成した圧縮ブロックのサイズが検知される。そこで、サイズリスト74として圧縮ブロックの使用バイト数を例えば1KB～16KB準備し、サイズごとに圧縮ブロックを分類した圧縮グループを組み合わせるための候補リストを作成する。

【0059】図10の場合、使用バイト数1KBには圧縮ブロック76-1、76-2が組合せ候補としてリンクされ、同様に使用バイト数2KB～5KBのそれぞれに同一サイズに分けられた圧縮ブロック76-3～76-8がリスト構造でリンクされている。このサイズリス

(11)

特開平8-249159

19

20

ト74を使用して圧縮グループを作成する手法は、

①同一サイズの圧縮ブロックを組み合わせる

②異なったサイズの圧縮ブロックを組み合わせる

の2つの方法がある。例えば①の方法は、使用バイト数1KBの圧縮ブロックの候補が16個できたときに、これらをまとめて1つの圧縮グループを作成する。使用バイト数が2KBの場合には、8つの圧縮ブロックをまとめればよい。

【0060】一方、②の方法は、サイズの異なる圧縮ブロックを組み合わせて16KBとなるようにする。この場合、同じサイズの圧縮ブロックを2つ以上使用してもよい。また圧縮グループを作成する他の方法としては、圧縮回路48が有するバッファメモリ（図4の圧縮データ格納用バッファメモリ132）に対する格納アドレスの順番に従って16KBサイズの圧縮グループを作ってもよい。また別の方法としては、圧縮回路48による時間的な圧縮の順番に従って16KBの圧縮グループを作ってもよい。

【0061】このような複数の圧縮ブロックの組合せによる16KBの圧縮グループの作成にあつては、図8の圧縮グループ70-nのように、4つの圧縮ブロック72-1~72-4を組み合わせて、余りを生ずることなく16KBとなることが望ましい。しかし、候補ブロックの組合せで16KBに満たない場合は、空き領域にダミーデータ（NULL）を付加して16KBの圧縮グループを作成してもよい。勿論、ディスク装置の実行的な使用率を高めるためには、可能な限り空き領域を生じないように圧縮グループを作ることが望ましい。

【0062】このようにして圧縮グループ作成部52で作成された圧縮グループは、圧縮キャッシュメモリ44に図8のように格納される。圧縮キャッシュメモリ44に格納された圧縮グループは、圧縮キャッシュメモリ44の使用容量が規定値に達すると、LRU部55により追出され、物理デバイスとしてのディスク14にライトバックされる。

【0063】ディスク装置14へのライトバックに際しては、ディスク装置14の論理セクタの空き状態を示しているフリーリンクドリスト56を参照して、リストの先頭から順番に空き論理セクタを割り当て、割り当てた論理セクタに圧縮グループを格納する。この圧縮グループに含まれる各圧縮ブロックの論理ブロック番号と格納先となる論理セクタ番号の対応関係は、マッピングテーブル54に登録される。

【0064】図11は本発明のマッピングテーブル54である。マッピングテーブル54には、ホストコンピュータ10で扱っている論理ブロック番号54-1に対応して圧縮ブロックを組み合わせた圧縮グループの格納先であるディスク装置14の論理セクタ番号54-2、圧縮グループの先頭から該当する圧縮ブロックまでの相対位置を示すオフセットセクタ数54-3、更に圧縮ブ

ックのサイズを示すセクタ長54-4が格納されている。

【0065】したがって、ホストコンピュータ10から通知された論理ブロック番号を用いて磁気ディスク制御装置12、具体的には図5のライトキャッシュ制御部62あるいはリードキャッシュ制御部64がディスク装置14から必要なデータを取り出す際には、マッピングテーブル54の参照で論理セクタ番号54-2を得ることでアクセスできる。またディスク装置14から読み出されて圧縮キャッシュメモリ44に展開された圧縮グループの中から必要とする論理ブロック番号の圧縮ブロックを取り出す際には、マッピングテーブル54のオフセットセクタ数54-3からグループ先頭からの位置を知り、セクタ長54-4分の圧縮ブロックを取り出し、これを伸長回路50で伸長して、非圧縮キャッシュメモリ42に伸長した論理ブロックを展開することができる。

【0066】図12は、論理グループに対する論理セクタの割当てに使用するフリーリンクドリストである。フリーリンクドリスト56は、論理セクタテーブル66の中の空き論理セクタ56-1、56-2、56-3をアドレス情報でつなげたリンク構造をもつ。まずフリーリンクドリストの空き論理セクタ56-1のアドレスは、先頭アドレスとしてMPU16が管理している。この先頭アドレスに対し、論理セクタテーブル66の先頭空き論理セクタ56-1に次の空き論理セクタ56-2のアドレスが格納され、以下同様に、後続する次の論理セクタのアドレスが次々と格納される。

【0067】このため、圧縮キャッシュメモリ44のライトバックの際には、MPU16がフリーリンクドリスト56の先頭アドレスをもつ空き論理セクタ56-1を、追出されてきた圧縮グループに割り当て、ディスク装置14への格納とマッピングテーブル54の作成を行う。図13は、非圧縮キャッシュメモリ42からディスク装置14に格納するまでの流れである。非圧縮キャッシュメモリ42は4KBのキャッシュブロック単位に管理されており、キャッシュ容量が規定値に達するとLRUによる追出しが行われ、論理ブロック単位に圧縮され、非圧縮キャッシュメモリ44上の16KBの圧縮グループ70-1~70-nに変換される。

【0068】圧縮キャッシュメモリ44についても、キャッシュ容量が規定値に達すると、LRUによる圧縮ブロックの追出しが行われ、ディスク装置14へ格納される。このとき、例えば圧縮キャッシュメモリ44から追出された4つの圧縮ブロック72-1~72-4の組合せでなる圧縮グループ70-nをそのままディスク装置14のフリーリンクドリスト56から割り当てられた論理セクタ80に格納してもよい。

【0069】更に、マッピングテーブル54よりセクタ情報78を作成し、セクタ情報78を圧縮グループ70-nの先頭に付加してディスク装置14の論理セクタ8

(12)

特開平8-249159

21

0に格納することが望ましい。ディスク装置14の論理セクタ80に格納する圧縮グループ70-nの先頭のセクタ情報78は、グループを構成している4つの圧縮ブロック72-1~72-4のそれぞれに関するセクタ情報78-1~78-4から構成されている。

【0070】例えば、圧縮ブロック72-1のセクタ情報78-1を例にとると、フラグ96-1、論理ブロック番号96-2およびオフセットセクタ数96-3で構成される。フラグ96-1は、セクタ情報が有効か否かを示す。フラグがセットされていればセクタ情報は有効である。論理ブロック番号96-2は、このセクタ情報78-1が示す圧縮ブロック72-1の論理ブロック番号である。オフセットアクセス数96-3は、圧縮グループ70-nの先頭からの圧縮ブロックまでの相対的な位置を示すセクタ数である。

【0071】このように、ディスク装置14に格納する圧縮グループの先頭に、グループに含まれる各圧縮ブロックの論理ブロック番号とグループ内の位置を示すオフセットアクセス数を含めておくことで、万が一、キャッシュメモリに設けているマッピングテーブル54が故障した場合には、ディスク装置14のセクタ情報を読み出すことでマッピングテーブル54を再構成することができる。

【0072】ここでセクタ情報78は、圧縮グループ70-nに最大で8つを格納可能とすると、1つの圧縮ブロックにつき64バイトであり、全体としては512バイトの固定長の領域とする。

### 3. データの記録処理

本発明の磁気ディスク制御装置12におけるデータの記録処理は、大きく分けて次の3つに分類される。

- 【0073】①キャッシュメモリにデータが存在せず、ディスク装置14にデータが存在する場合
  - ②非圧縮キャッシュメモリ42にデータが存在する場合
  - ③圧縮キャッシュメモリ44にデータが存在する場合
- そこで①~③のそれぞれについて、磁気ディスク制御装置12の処理動作を説明する。

【0074】図14は、ホストコンピュータ10が要求した論理ブロック番号に該当するデータがキャッシュメモリに存在しなかった場合、即ち非圧縮キャッシュメモリ42にも存在せず圧縮キャッシュメモリ44にも存在しなかった場合の処理である。まずホストコンピュータ10からの論理ブロック番号を用いた書込要求がライトキャッシュ制御部62に行われ、ライトキャッシュ制御部62はヒット判定回路30にヒット判定を依頼する。

【0075】ヒット判定回路30は、ホストコンピュータ10からの論理ブロック番号を使用してキャッシュテーブル40を参照する。この場合、書込要求対象となった論理ブロック番号は、非圧縮キャッシュメモリ42および圧縮キャッシュメモリ44のいずれにも存在しないことからミスヒットとなる。ヒット判定回路30からの

22

ミスヒットの判定結果を受けたライトキャッシュ制御部62は、ホストコンピュータ10からの書込論理ブロックについて、所定のハッシングによりキャッシュ内のアドレスを求め、非圧縮キャッシュメモリ42にキャッシュブロック68として格納し、ホストコンピュータ10にデバイスエンドを通知する。この非圧縮キャッシュメモリ42の格納と同時に、キャッシュテーブル40に対し図7のような論理ブロック番号40-1に対応したキャッシュ内のアドレス40-2、データ長40-3の登録が行われ、この場合、非圧縮データであることから、圧縮フラグ40-4はオフにリセットされており、また更新フラグ40-5も新規データの格納であることからオフにリセットされている。

【0076】図15は、非圧縮キャッシュメモリ42にデータが存在した場合の処理動作である。ホストコンピュータ10からの論理ブロック番号を指定した書込要求を受けたライトキャッシュ制御部62は、ヒット判定回路30に判定処理を依頼する。ヒット判定回路30はキャッシュテーブル40を参照し、該当する論理ブロック番号を検索し、その圧縮フラグがオフにリセットされていることから、非圧縮キャッシュメモリ42に存在しているヒット判定を行う。

【0077】このヒット判定を受けてライトキャッシュ制御部62は、ホストコンピュータ10からの書込論理ブロックを使用して、ヒット判定で特定された非圧縮キャッシュメモリ42内のキャッシュブロック68を更新する。キャッシュブロック68の更新が済むと、キャッシュテーブル40の更新フラグ40-5（図7参照）をチェックする。もし更新フラグがオフにリセットされていれば更新フラグをオンにセットする。

【0078】更新フラグをオンにセットした場合には、圧縮キャッシュメモリ44の中に、更新したキャッシュブロック68の伸長回路50による展開に使用した圧縮グループ70が存在していることから、キャッシュブロック68の更新により圧縮グループ70を再構成しなければならない。そこで、マッピングテーブル54の中から、更新が行われたキャッシュブロック68を含んでいる圧縮グループ70に割り当てた論理セクタ番号を消去する。

【0079】具体的には、マッピングテーブル54から圧縮グループ70を構成する各圧縮ブロックに対応する論理ブロック番号を全て消去する。そして、更新したキャッシュブロック68以外の圧縮ブロック70を構成している圧縮ブロックを圧縮グループ作成部52に渡し、新たな圧縮グループを作成するための候補リストに加える。

【0080】このように、圧縮グループを構成する特定の圧縮ブロックの伸長ブロックが更新された場合には、圧縮グループを解散し、更新対象以外の圧縮ブロックについては、再度、圧縮グループを作成する候補に戻すこ

(13)

特開平8-249159

23

とで、全く新しい組合せの圧縮グループを作成してディスク装置14に格納するようになる。図16は、圧縮キャッシュメモリ44にデータが存在した場合である。ホストコンピュータ10からの論理ブロック番号の指定による書込要求を受けてライトキャッシュ制御部62がヒット判定回路30にヒット判定を依頼すると、キャッシュテーブル40の参照により、圧縮キャッシュメモリ44に該当する論理ブロック番号の圧縮ブロック72が存在していることが判る。

【0081】そこでライトキャッシュ制御部62は、マッピングテーブル54を参照して、圧縮グループ70に含まれている圧縮ブロック72の先頭からの相対位置とデータ長を認識し、圧縮ブロック72を取り出して伸長回路50で伸長して、非圧縮キャッシュメモリ42にキャッシュブロック68として展開する。続いてライトキャッシュ制御部62は、ホストコンピュータ10からの書込要求ブロックによって、非圧縮キャッシュメモリ42に展開したキャッシュブロック68を更新する。

【0082】キャッシュブロック68の更新が済むと、キャッシュテーブル40の該当する更新フラグをオンにセットする。またマッピングテーブル54に対し、圧縮ブロック72を取り出した圧縮グループ70に含まれる他の圧縮ブロックの論理ブロック番号を全て消去し、また圧縮グループ70に割り当てていた論理セクタも消去し、フリーリンクドリスト56に戻す。マッピングテーブル54の消去が済むと、更新が行われた圧縮ブロック72以外の圧縮グループ70に含まれる他の圧縮ブロックを圧縮グループ作成部52に戻し、新たな圧縮グループを組み合わせるための候補とする。

【0083】図17は、非圧縮キャッシュメモリ42の使用率が高まってライトバックのためにキャッシュブロックの追出しが行われたときの圧縮処理と、同様に圧縮キャッシュメモリ44のキャッシュ容量の使用率が高まって圧縮グループをライトバックするためのキャッシュ追出しが行われたときの処理である。まず非圧縮キャッシュメモリ42において、LRU部46で管理されているキャッシュブロックのうち、キャッシュ使用率の増加に伴ってキャッシュブロック68-1~68-mがディスク装置14よりのライトバックのために追い出されたとする。この非圧縮キャッシュメモリ42から追い出されたキャッシュブロック68-1~68-mは、圧縮回路48に供給され、ブロック単位に圧縮され、圧縮グループ作成部52で16KB単位の圧縮グループにまとめられて、圧縮キャッシュメモリ44に格納される。

【0084】このように非圧縮キャッシュメモリ42から追い出されて圧縮されたキャッシュブロック68-1~68-mに関するキャッシュテーブル40の内容は、図7に示すキャッシュブロック68-1~68-mに対応する論理ブロック番号40-1の中の圧縮フラグ40-4がセットされてオンとなる。これによってキャッシ

24

ュブロック68-1~68-mが圧縮により圧縮キャッシュメモリ44に存在することが判る。

【0085】圧縮キャッシュメモリ44に格納されている複数の圧縮データのうち、キャッシュ容量の使用率が高まって、例えばLRU部55により圧縮グループ70の追出しが行われたとする。この圧縮グループ70には、非圧縮キャッシュメモリ42から追い出されたキャッシュブロックの圧縮で得られた複数の圧縮ブロック72-1~72-mが組み合わされている。

【0086】このときライトキャッシュ制御部62は、フリーリンクドリスト56を参照し、MPU16で管理している先頭アドレスの空き状態にある論理セクタを追い出された圧縮グループ70に割り当て、デバイス書込部58によりディスク装置14に書き込まれることになる。ここで、ライトキャッシュ制御部62によるディスク装置14への書込みに際しては、フリーリンクドリスト56で空き状態にある論理セクタの先頭位置を割り当てているが、この論理セクタの割当てをハッシングにより行うようにしてもよい。しかしながら、ハッシングによる論理セクタの割当てでは圧縮グループがランダムにディスク装置に格納され、シーク性能が低下する恐れがある。

【0087】そこで、論理セクタをディスク装置のシリンダアドレスごとに組み分けしておき、同一シリンダアドレスに含まれる論理セクタ内でハッシングにより圧縮グループに割り当てることで、圧縮グループのライトバックの際に同一シリンダ位置での格納が行われ、1回シークすれば、その後はヘッド切替だけで済むことから、圧縮データのライトバック時のアクセス性能を高めることができる。

【0088】図18は、本発明の磁気ディスク制御装置における書込処理のフローチャートである。まずステップS1で、ホストコンピュータ10から論理ブロックアドレス（論理ブロック番号）とデータ（論理ブロック）を受診すると、ステップS2で、キャッシュメモリのヒット判定を行う。非圧縮キャッシュメモリ42および圧縮キャッシュメモリ44のいずれにもデータが存在しないときには、ステップS8~S14の処理を行う。

【0089】これは図14に示した非圧縮キャッシュメモリ42に対するキャッシュブロック68のステージングである。即ち、ステップS8で、非圧縮キャッシュメモリ42に新規データを記録する。続いてステップS9でマッピングテーブル54を参照し、ステップS10で、マッピングテーブルに登録しているか否かチェックする。マッピングテーブル54の登録がなければ新規データであることから、ステップS14で、キャッシュテーブルに新規データの記録に関する事項を登録して更新する。

【0090】マッピングテーブル54に登録があれば、ディスク装置14に格納された圧縮グループの中に圧縮

(14)

特開平8-249159

25

ブロックとして含まれていることから、ステップS11で、論理ブロック番号に対応する論理セクタ番号を調べ、ディスク装置14から該当する圧縮グループを読み出し、圧縮キャッシュメモリ44に転送する。圧縮キャッシュメモリ44に転送した圧縮グループの中から、マッピングテーブル54で得られた圧縮ブロックを取り出し、伸長回路50で伸長して、非圧縮キャッシュメモリ42に展開し、展開した論理ブロックを新規データにより更新する。更新が済むと、ステップS13のマッピングテーブル更新処理に進む。

【0091】このマッピングテーブル更新処理は、図19のフローチャートのようになる。まずステップS14で、キャッシュテーブル40を参照し、該当する論理ブロック番号の更新フラグをオンにセットする。次にステップS14で、マッピングテーブル54から更新された論理セクタを削除する。この論理セクタの削除により、圧縮キャッシュメモリ44にディスク装置14から読み出された圧縮データの中の更新された圧縮ブロック以外の他の圧縮ブロックは圧縮グループ作成部52に戻され、新たな圧縮グループの作成候補となる。続いてステップS16で、削除した論理セクタ番号をフリーリンクドリスト56に登録し、図18のメインルーチンに戻って一連の処理を終了する。

【0092】尚、ステップS11～S13の更新処理を行わず、非圧縮キャッシュメモリ42空野LRUによるデータ追い出しに伴うライトバック処理としてディスク装置14のデータを更新してもよい。次に、非圧縮キャッシュメモリ42でキャッシュヒットとなった場合には、ステップS3、S4、S5の処理となる。これは図15の処理動作である。まずステップS3で、非圧縮キャッシュメモリ42のヒットしたアドレスの該当するキャッシュブロック68をホストコンピュータ10からの新規データで更新する。続いてキャッシュテーブル40の該当するキャッシュブロック番号の更新フラグがオンしているか否かをチェックする。もし更新フラグが既にオンしていれば2回目以降の更新であることから、マッピングテーブル54からのセクタ番号の消去は不要であり、処理を終了する。

【0093】キャッシュテーブル40の論理ブロック番号に対応する更新フラグがオフにリセットされていた場合には、ステップS5のマッピングテーブル更新処理に進む。マッピングテーブル更新処理は、図19のサブルーチンに示すように、ステップS14で、キャッシュテーブル40の論理ブロック番号に対応する更新フラグをオンにセットし、ステップS16で、マッピングテーブル54から更新された論理セクタを削除して圧縮グループを解散し、更新した論理ブロック以外の論理ブロックに対応する圧縮ブロックを圧縮グループ作成部52に戻して、新たなグループ作成の候補とする。そしてステップS16で、削除した論理セクタ番号をフリーリンクド

26

リスト56に戻し、再び図7のメインルーチンに戻って一連の処理を終了する。

【0094】更に、圧縮キャッシュメモリ44でキャッシュヒットとなった場合には、ステップS6、S7、S5の処理が行われる。この処理は、図16に示したとおりである。まず、圧縮キャッシュメモリ44に該当する論理ブロックの圧縮ブロック72を含む圧縮グループ70が存在していることから、ステップS6で、マッピングテーブル54を参照して圧縮ブロック72を取り出し、伸長回路50により伸長して、非圧縮キャッシュメモリ42にキャッシュブロック68として転送し、転送後に、ホストコンピュータ10からの書込ブロックによってデータを更新する。

【0095】次にステップS7で、更新した論理ブロック番号の更新フラグがオンかオフかをチェックする。既に更新フラグがオンにセットされていればマッピングテーブル54の消去が済んでいることから、処理を終了する。更新フラグがオフにリセットされていた場合には、初めての更新であることから、ステップS5のマッピングテーブル更新処理を行う。

【0096】即ち、図19のステップS14に示すように、キャッシュテーブル40の更新フラグをオンにセットし、ステップS15で、マッピングテーブル54から更新された論理ブロックの圧縮ブロックを含む論理セクタを削除する。これによって、圧縮キャッシュメモリ44の更新対象となる圧縮ブロック72を含む圧縮グループ70が解散され、圧縮ブロック72以外の圧縮ブロックは圧縮グループ作成部52に戻されて、新たなグループを構成する候補となる。続いてステップS16で、削除した論理セクタ番号をフリーリンクドリスト56に登録し、図18のメインルーチンに戻って一連の処理を終了する。

#### 4. データの読出処理

図20は本発明の磁気ディスク制御装置12におけるリード処理であり、要求データが非圧縮キャッシュメモリ42および圧縮キャッシュメモリ44に存在せず、ディスク装置14に存在していた場合を示している。

【0097】ホストコンピュータ10からの論理ブロック番号を指定した読出要求をリードキャッシュ制御部64で受けると、ヒット判定回路30にヒット判定を依頼する。ヒット判定回路30は、論理ブロック番号によりキャッシュテーブル40を参照するが、この場合には、非圧縮キャッシュメモリ42、圧縮キャッシュメモリ44の両方に存在しないことから、キャッシュテーブルに指定された論理番号の登録はなく、ミスヒットを判定する。

【0098】ヒット判定回路30よりミスヒットの判定結果が得られると、リードキャッシュ制御部64は指定された論理ブロック番号によりマッピングテーブル54を参照し、論理セクタ番号を認識して、デバイス読出部

(15)

特開平 8-249159

27

60によりディスク装置14から検索された論理セクタに格納されている圧縮グループ70を読み出して、圧縮キャッシュメモリ44に展開する。

【0099】続いてマッピングテーブル54から論理ブロック番号に対応するグループ内の先頭からの相対位置およびデータ長を認識して、圧縮ブロック72を取り出し、伸長回路50により伸長して、非圧縮キャッシュメモリ42に展開する。そして、非圧縮メモリ42に展開した論理ブロックをホストコンピュータ10に転送して一連の処理を終了する。このディスク装置14にデータが存在するときの処理は、図21のフローチャートのステップS1、S2、S10～S12、S6～S9の処理になる。

【0100】まずステップS1で、ホストコンピュータ10からの読出要求による論理ブロック番号を受診し、ステップS2で非圧縮キャッシュメモリ42および圧縮キャッシュメモリ44のいずれにも存在しないことを認識すると、ステップS10に進む。ステップS10では、マッピングテーブル54を用いて論理ブロック番号から論理セクタ番号を導き、ステップS11で、該当する論理セクタの圧縮グループをディスク装置14から読み出して、圧縮キャッシュメモリ44に転送する。

【0101】続いてステップS14で、ホストコンピュータ10から要求されたブロックに該当する圧縮ブロックを圧縮キャッシュメモリ44に転送された圧縮グループ30の中から取り出して、伸長回路50に転送して復元し、ステップS6で、非圧縮キャッシュメモリ42にデータを転送する。次にステップS7で、転送したデータを非圧縮キャッシュメモリから読み出し、ステップS8でホストコンピュータに転送する。最終的に、ステップS9で、キャッシュテーブル40に非圧縮キャッシュメモリ42に復元されたキャッシュブロック68が存在し、圧縮キャッシュメモリ44に読出対象となった論理ブロックの圧縮ブロック72と同じグループの圧縮グループ70に含まれる他の圧縮ブロックが存在することを各論理ブロック番号と共に登録する。なお、読出要求を受けた論理ブロック番号については、圧縮キャッシュメモリ44に圧縮ブロック72が存在し、非圧縮キャッシュメモリ42に復元された圧縮ブロック68も存在するが、この場合には圧縮フラグはオフにリセットし、非圧縮キャッシュメモリ48にキャッシュブロック68が存在することのみをキャッシュテーブル40に登録しておくべき。

【0102】これに対し、非圧縮キャッシュメモリ42に復元されていない圧縮グループ70に含まれる他の論理ブロックについては、圧縮キャッシュメモリ44に存在することを示すために圧縮フラグを1にセットしておく。次に、ホストコンピュータ10からのリード要求に対し、非圧縮キャッシュメモリ40にデータが存在した場合を説明する。この場合の処理は、図21のステップ

28

S3、S4の処理になる。即ち、ステップS3で、キャッシュテーブル40の参照で得られた非圧縮キャッシュメモリ42のアドレスからデータを読み出して、ステップS4で、ホストコンピュータ10に転送して処理を終了する。

【0103】更に、ホストコンピュータ10からの読出要求に対し、データが圧縮キャッシュメモリ44に存在した場合は、図21のステップS5～S9の処理となる。まずステップS5で、マッピングテーブル54を参照し、読出要求が行われた論理ブロックの圧縮キャッシュメモリ44に格納されている圧縮グループ70内での先頭からの相対位置とデータ長を認識して、該当する圧縮ブロック72を取り出し、伸長回路50で復元して、ステップS6で非圧縮キャッシュメモリ42に転送する。

【0104】続いてステップS7で、非圧縮キャッシュメモリ42からデータを読み出し、ステップS8で、該当するブロックをホストコンピュータ10に転送する。続いてステップS9で、非圧縮キャッシュメモリ42にデータを復元したことから、読出要求を行った論理ブロック番号の圧縮フラグを0にリセットするキャッシュテーブル40の更新を行う。

## 5. 二重化構成

図22は本発明の他の実施例であり、図2の磁気ディスク制御装置12を2系統設けて、ホストコンピュータおよびディスク装置に対し二重化構成としたことを特徴とする。

【0105】まず、システム1は磁気ディスク制御装置12-1で構成し、システム2は磁気ディスク制御装置12-2で構成する。磁気ディスク制御装置12-1、12-2は同じ構成をもち、MPU16-1、16-2、ROM18-1、18-2、RAM20-1、20-2、ホストインタフェース制御装置22-1、22-2、ディスクインタフェース制御装置24-1、24-2、キャッシュ制御部26-1、26-2、キャッシュメモリ28-1、28-2、ヒット判定回路30-1、30-2、圧縮伸長回路32-1、32-2、内部バス34-1、34-2を備える。

【0106】ホストコンピュータ10に対しシステム1、2の磁気ディスク制御装置12-1、12-2は、独立したチャネルバス36-1、36-2で接続される。またディスクアレイで構成されるディスク装置14-1、14-6に対しても、独立したデバイスバス38-1、38-2で接続される。磁気ディスク制御装置12-1の内部バス34-1と、磁気ディスク制御装置12-2の内部バス34-2は、ブリッジバッファIF152および周辺ロジック154を介して接続され、それぞれのMPU16-1、16-2間でコマンドメッセージおよびデータのやり取りができる。

【0107】このような二重化構成は、ハードウェア的

(15)

特開平8-249159

29

30

には、図3に示したシステム1、システム2のモジュール構成に対応する。二重化により、システム1、システム2の構成する磁気ディスク制御装置12-1、12-2の機能は、図5の機能ブロックに示したと基本的に同じである。即ち、磁気ディスク制御装置12-1、12-2は共にライトキャッシュ制御部62およびリードキャッシュ制御部64としての機能をもち、ホストコンピュータ10からの要求に対し、基本的に独立した動作を行うが、例えばホストコンピュータ10からの要求を受けたシステム1でミスヒットとなった場合には、直ちにディスク装置をアクセスせず、システム2のキャッシュヒットを判定する。またシステム1のキャッシュメモリ28-1でデータの更新が行われた場合には、システム2のキャッシュメモリ28-2との同一性を保証するため、更新データの複写とマッピングテーブルの更新が行われることになる。

#### 6. 二重化システムの書込処理

図23は、ホストコンピュータ10からシステム1の磁気ディスク制御装置12-1が書込要求を受けたときの処理である。まずステップS1で、ホストコンピュータ10から論理ブロック番号とデータ（論理ブロック）を受信すると、ステップS2で、システム1のキャッシュメモリ28-1についてヒット判定を行う。キャッシュメモリ28-1のヒット判定は、図5の機能ブロックに示したように、非圧縮キャッシュメモリ42と圧縮キャッシュメモリ44について行う。

【0108】システム1の非圧縮キャッシュメモリ42でヒットすると、ステップS4の更新処理1を行う。システム1の圧縮キャッシュメモリ44でヒットすると、ステップS5の更新処理2を行う。システム1の非圧縮キャッシュメモリ42および圧縮キャッシュメモリ44のいずれもミスヒットとなった場合には、ステップS3に進み、システム2のディスク制御装置12-2に対しヒット判定を依頼する。

【0109】これにより、システム2の非圧縮キャッシュメモリ42でヒットした場合には、ステップS6の更新処理3を行う。またシステム2の圧縮キャッシュメモリ44でヒットした場合には、ステップS7の更新処理4を行う。更に、システム1およびシステム2のいずれのキャッシュメモリ28-1、28-2でもミスヒットとなった場合には、ステップS8の更新処理5を行う。

【0110】図24は、図23のステップS2でシステム1の非圧縮キャッシュメモリ42でヒットした場合の更新処理1と、圧縮メモリ44でヒットした場合の更新処理2である。更新処理1を図5の機能ブロックについて説明すると次のようになる。まずステップS1で、非圧縮キャッシュメモリ42に書込要求を行った論理ブロック番号のデータが存在することから、このデータを書込データで更新する。次にステップS2で、キャッシュテーブル40の更新したデータのブロック番号に対応す

る更新フラグを1にセットする。

【0111】次にステップS3で、更新フラグが既に1にセットされていたか否かチェックする。既に1にセットされていれば、論理セクタのマッピングテーブル54からの消去は必要ないため、処理を終了する。更新フラグが更新前に0にリセットされておりステップS1で更新フラグを1にセットした場合には、初めての更新であることから、更新ブロックをシステム2に転送して、システム2の非圧縮キャッシュメモリ42に複写する。続いてステップS5で、システム2のキャッシュテーブル40の該当する論理ブロック番号の更新フラグを1にセットする。続いてステップS6で、マッピングテーブル更新処理に進む。

【0112】このマッピングテーブル更新処理は、図25のステップS11、S12となる。まずステップS11で、マッピングテーブル44から、更新された論理ブロックの圧縮ブロックが含まれている圧縮グループに割り当てた論理セクタ番号を削除する。このとき圧縮キャッシュメモリ44に非圧縮キャッシュメモリ42で更新した論理ブロックを伸長した圧縮グループが残っている場合には、この圧縮グループに含まれる他の論理ブロックの圧縮ブロックを圧縮グループ作成部52に戻して、新たな圧縮グループを作成するための候補とする。次にステップS12で、削除した論理セクタ番号をフリーリンクドリスト56に戻す。

【0113】次に図24の更新処理2を説明する。更新処理2は、図23のように、システム1の圧縮キャッシュメモリ44でヒットした場合の処理である。まずステップS1で、ホストコンピュータ10から受信した書込対象となる論理ブロックを非圧縮キャッシュメモリ42に記録する。次にステップS2で、更新されるべき論理ブロック番号によりマッピングテーブル54を参照し、圧縮キャッシュメモリ44に存在している該当する圧縮グループのグループ先頭位置からのオフセットセクタ数とデータ長から該当する圧縮ブロックを取り出し、伸長回路50で復元して非圧縮キャッシュメモリ42に転送し、ステップS1で記録した論理ブロックによるデータ更新を行う。

【0114】非圧縮キャッシュメモリ42上でのデータ更新が済むと、更新処理1のステップS2～S6と同じ処理を行う。即ち、ステップS2でシステム1のキャッシュテーブル40の論理ブロック番号に対応する更新フラグを1にセットする。次にステップS3で、更新フラグが既に1にセットされていたか否かチェックし、既に1にセットされていれば処理を終了する。

【0115】更新フラグが0にリセットされていた場合には、最初の更新であることから、ステップS4に進み、システム2の非圧縮キャッシュメモリに更新ブロックを複写し、ステップS5でシステム2のキャッシュテーブルの該当する更新フラグを1にセットし、ステップ

(17)

特開平8-249159

31

S6のマッピングテーブル更新処理を行う。即ち、図25のステップS11で、マッピングテーブル54から更新ブロックに対応する論理セクタを削除し、ステップS12で、削除した論理セクタ番号をフリーリンクドリスト56に登録する。

【0116】図26は、更新処理3と更新処理4を示す。まず更新処理3は、図23のように、システム1でキャッシュミスヒットとなり、システム2の非圧縮キャッシュメモリ42でヒットした場合の処理である。まずステップS1で、ホストコンピュータ10からの論理ブロックをシステム1の非圧縮キャッシュメモリ42に記録する。

【0117】次にステップS2で、システム2のキャッシュヒットとなった該当ブロックを非圧縮キャッシュメモリ42から読み出して、システム1の非圧縮キャッシュメモリ42に複写し、ステップS2で記録した論理ブロックにより複写ブロックを更新し、更にシステム1のキャッシュテーブル40の更新フラグを1にセットする。続いてステップS3の二重化処理に入る。

【0118】図27は、二重化処理の詳細である。まずステップS11で、システム1の非圧縮キャッシュメモリ42で更新された新データをシステム2の非圧縮キャッシュメモリ42に複写し、システム2のキャッシュテーブル40の該当する更新フラグを1にセットする。続いてステップS4で、マッピングテーブル更新処理を行う。

【0119】即ち、図25のように、ステップS11で、システム2のマッピングテーブル56から更新された論理ブロックの論理ブロック番号に対応する論理セクタ番号を検索して、論理セクタ番号を削除する。このとき同時に更新ブロックと同じ圧縮グループに含まれている他の論理ブロックもマッピングテーブル56から削除され、圧縮グループ作成部52に戻されて、新たなグループ作成の候補に戻る。

【0120】なお、キャッシュメモリ上に該当ブロックが存在しないシステム1側についても、マッピングテーブル54の同一性を保証するため、更新ブロック番号に対応する論理セクタ番号の削除を行なっていくことが望ましい。次に、図26の更新処理4を説明する。更新処理4は、図23のように、システム1のキャッシュメモリでミスヒットとなりシステム2の圧縮キャッシュメモリ44でヒットした場合の処理である。まずステップS1で、システム1の非圧縮キャッシュメモリ42にホストコンピュータ10からの該当する論理ブロックを記録する。

【0121】次にステップS2で、システム2の圧縮キャッシュメモリ44の該当する圧縮データの中から、マッピングテーブル54の参照で得られた圧縮ブロックを取り出して伸長回路50で復元して、非圧縮キャッシュメモリ42に転送し、更にシステム1の非圧縮キャッシュ

32

メモリ42に転送して、ステップS1で記録した論理ブロックによるデータ更新を実行する。

【0122】データ更新が済んだならば、システム1のキャッシュテーブル40の該当する論理ブロック番号の更新フラグを1にセットする。次にステップS3で、更新データをシステム2の非圧縮キャッシュメモリ42に転送し、同様にシステム2のキャッシュテーブル40を更新する。続いてステップS4のマッピングテーブル更新処理を行う。

【0123】即ち、図25のステップS11で、システム2のマッピングテーブル54から更新ブロック番号に対応する論理セクタ番号を削除する。これによって、更新ブロック番号と同じ圧縮グループに含まれていた他の論理ブロック番号の圧縮ブロックは圧縮グループ作成部52に戻され、新たなグループ形成の候補となる。次にステップS14で、削除した論理セクタ番号をフリーリンクドリスト56に登録する。この場合も、システム1側のマッピングテーブル54の消去を同様に行って、同一性を保証することが望ましい。

【0124】図28は更新処理5であり、図23のように、システム1およびシステム2の両方でミスヒットとなってディスク装置14からデータを読み出して更新する処理である。まずステップS1で、ホストコンピュータ10からの書込要求を受けた論理ブロックをシステム1の非圧縮キャッシュメモリ42に記録する。次にステップS2で、システム1のマッピングテーブル54を検索する。続いてステップS3で、要求された論理ブロック番号がマッピングテーブル54に登録されているか否かチェックする。

【0125】もし登録されていなければ、新規データの書込みであることから、ステップS7で、キャッシュテーブル40に新たな論理ブロック番号と対応するキャッシュ内のアドレスおよびデータ長を登録し、この場合、圧縮フラグは0にリセット、更新フラグも0にリセットとするテーブル更新を行う。一方、マッピングテーブル54に要求された論理ブロック番号の登録があった場合には、ディスク装置14に格納されていることから、論理セクタ番号を調べ、ディスク装置14から該当する論理セクタの圧縮グループを読み出し、圧縮キャッシュメモリ44に転送する。次にステップS5で、圧縮キャッシュメモリ44に転送された圧縮グループの中から、マッピングテーブル54の参照で得られたグループ先頭位置からの相対位置およびデータ長に基づき、該当する圧縮ブロックを取り出し、伸長回路50で復元して非圧縮キャッシュメモリ42に転送して、ステップS1で記録した論理ブロックによるデータ更新を実行する。

【0126】更新が済んだならば、キャッシュテーブル40の該当する論理ブロック番号の更新フラグを1にセットする。続いて図27のステップS11～S14の二重化処理を行う。即ち、ステップS11で、更新が済ん

(18)

33

だブロックをシステム2の非圧縮キャッシュメモリ42に複写して、システム2のキャッシュテーブル40を同様に更新する。

【0127】続いて図25のマッピングテーブル更新処理に入り、ステップS11で、システム1のマッピングテーブル54から更新ブロックの論理ブロック番号に対応する論理セクタ番号を削除する。これによって圧縮キャッシュメモリ44の更新ブロック以外の圧縮ブロックは圧縮グループ作成部52に属され、新たなグループ作成候補となる。

【0128】最終的にステップS12で、削除した論理セクタ番号をフリーリンクドリスト56に登録する。またシステム2のマッピングテーブル54についても、ディスク装置14に対する更新した論理ブロック番号と論理セクタ番号の対応の登録があることから、同様にして論理セクタ番号を削除し、フリーリンクドリスト56に登録させる。

【0129】なお上記の二重化構成における書込処理は、システム1に対するホストコンピュータからの書込要求に対する処理を例にとっているが、システム2に対しホストコンピュータからの書込要求が行われた場合は、システム2側をメインとした同様の処理を行うことになる。

#### 7. 二重化構成の読出処理

図29のフローチャートは、図22の二重化構成における読出処理である。この処理を図5の機能ブロックについて説明すると次のようになる。

【0130】まずシステム1に対し、ホストコンピュータ10より読出要求が行われると、ステップS1で論理ブロック番号が受信される。この論理ブロック番号によるヒット判定により、ステップS2でシステム1の非圧縮キャッシュメモリ42のヒットが判定されると、ステップS3で、該当ブロックをホストコンピュータ10に転送して処理を終了する。

【0131】またステップS2で、システム1の圧縮キャッシュメモリ44でヒットした場合には、ステップS4で、論理ブロック番号によるマッピングテーブル54の参照で圧縮キャッシュメモリ44の圧縮グループの中の該当する圧縮ブロックを取り出し、伸長回路50で復元して、非圧縮キャッシュメモリ42に転送する。続いてステップS5で、該当ブロックをホストコンピュータ10に転送し、最終的にステップS6で、キャッシュテーブル40の該当する論理ブロック番号の圧縮フラグを0にリセットし、非圧縮データとして非圧縮キャッシュメモリ42にデータが存在することを示すテーブル更新を行う。

【0132】一方、システム1の非圧縮キャッシュメモリ42および圧縮キャッシュメモリ44のいずれでもミスヒットとなった場合には、ステップS7に進み、システム2のキャッシュメモリのヒット判定を行う。このシ

特開平8-249159

34

ステム2のヒット判定でシステム2の非圧縮キャッシュメモリ42のヒットとなった場合には、ステップS8で、該当ブロックを読み出してホストコンピュータ10に転送する。このシステム2からの該当ブロックの転送は、システム1を経由して行ってもよいし、直接、システム2からホストコンピュータ10に転送してもよい。

【0133】またシステム2の圧縮キャッシュメモリ44のヒットとなった場合には、ステップS9で、要求された論理ブロック番号によるマッピングテーブル54の参照で圧縮キャッシュメモリ44の圧縮データの中の該当する圧縮ブロックを取り出して伸長回路50で復元し、非圧縮キャッシュメモリ42に転送する。続いてステップS10で、非圧縮キャッシュメモリ42に転送したブロックをホストコンピュータ10に転送する。この場合の転送も、システム2から直接行ってもよいしシステム1を経由して行ってもよい。最終的に、ステップS11で、システム2のキャッシュテーブル40の該当する論理ブロック番号の圧縮フラグを0にリセットし、非圧縮キャッシュメモリ42にデータが存在することを示すテーブル更新を行う。

【0134】更に、システム1およびシステム2の両方でミスヒットとなった場合には、システム1において要求された論理ブロック番号によりマッピングテーブル54をステップS12で検索する。次にステップS13で、マッピングテーブル56の検索で得られた論理セクタ番号の圧縮グループをディスク装置から読み出し、圧縮キャッシュメモリ44に転送する。

【0135】続いてステップS14で、読み出した圧縮グループの中の該当する圧縮ブロックをマッピングテーブル54のオフセットセクタ数およびデータ長から取り出して伸長回路50で復元し、非圧縮キャッシュメモリ42に転送する。次にステップS15で、該当するブロックを非圧縮キャッシュメモリ42からホストコンピュータに転送し、最終的に、ステップS16で、システム1のキャッシュテーブル40に、新たに読み出した論理ブロック番号、キャッシュ内アドレスを登録し、このときの更新フラグは0、圧縮フラグは0にリセットさせるテーブル更新を行う。

8. 位置フラグの付加による圧縮データのブロック化  
上記の実施例にあっては、図9のように、4KBの論理ブロック68を圧縮して得た圧縮データがディスク装置の最小物理セクタの整数倍に満たない場合は、ダミーデータ82を付加して物理最小セクタの整数倍のサイズをもつ圧縮ブロックを作っている。

【0136】これに対し本発明の他の実施例として、圧縮データにダミーデータを付加せずに連続的に組み合わせる16KBの圧縮グループを作り、圧縮データのグループ内での位置を特定するため、圧縮データの先頭に位置フラグを設けるようにしてもよい。図30は、ダミーデータの付加を必要としない場合の圧縮グループ形成の

(19)

特開平 8-249159

35

説明である。今、非圧縮キャッシュメモリ 42 には 4 K B のサイズをもつキャッシュブロック 68-1~68-8 として論理ブロック番号 #50~#57 が格納されていたとする。これら 8 つのキャッシュブロック 68-1~68-8 が LRU による追出しを受けて圧縮回路 48 で圧縮され、圧縮キャッシュメモリ 44 の 16 K B の圧縮グループ 70-1、70-2 に示すように、圧縮データを連続的に格納してグループを作成している。

【0137】図 31 は図 30 の圧縮グループ 70-1 を取り出しており、論理ブロック番号 #50~#55 の 6 つの圧縮データ 72-1~72-6 が組み合わされている。圧縮データ 72-1~72-6 のそれぞれは、先頭の圧縮データ 72-1 を取り出して示すように、先頭に位置フラグ 90 を付加しており、その後ろに圧縮データ 92 が存在する。

【0138】位置フラグ 90 は、圧縮グループ 70-1 内における圧縮データ 92 の位置を表わす情報であり、この位置フラグ 90 は、基本的にはフラグ 1、フラグ 2、フラグ 3 の 3 種類で構成される。フラグ 1 は、グループの先頭位置にあることを示す。フラグ 2 は、グループの途中にあることを示す。更にフラグ 3 は、グループの終端から次のグループの先頭にあることを示す。

【0139】図 30 の圧縮キャッシュメモリ 44 に格納されている圧縮グループ 70-1、70-2 が LRU に従って追い出されると、図 5 の機能ブロックと同様に、フリーリンクドリストに従った論理セクタの割付けが行われ、ディスク装置 14 への格納が行われる。このとき図 32 に示すようなマッピングテーブル 54 が作成される。

【0140】マッピングテーブル 54 は、論理ブロック番号 51 に対応して割り当てられた論理セクタ番号 54-2 が登録され、更に圧縮データ 92 の先頭に設けた位置フラグ 54-5 が登録される。図 32 のマッピングテーブル 54 は、図 30 の圧縮キャッシュメモリ 44 の圧縮グループ 70-1、70-2 をディスク装置に格納したときの内容となる。

【0141】即ち、論理ブロック番号 #50~#55 で 1 つの圧縮グループを構成して、論理セクタ番号 10 番に格納されていることが判る。また論理ブロック番号 #55 の論理セクタ番号は 10 番であるが、位置フラグがフラグ 3 であることから、論理セクタ番号 10 の終端から次の論理セクタ番号 11 の先頭にかけて格納されていることが判る。

【0142】更に、位置フラグはグループ内における相対的な位置を表わすことから、グループの途中であることを示すフラグ 2 については、論理ブロック番号 #52~#54 に示すように、2-1~2-4 と、途中に存在する位置の順番を示すフラグビットが付加されている。位置フラグのビット数は、16 K B の圧縮グループに格納する圧縮ブロックの最大数を例えば 8 に制限している

36

ことから、数ビット程度でよい。

【0143】図 33 は、論理ブロック単位に圧縮された圧縮データの先頭に位置フラグを付加して圧縮グループを作成した場合の言込処理であり、図 5 の機能ブロックを参照して説明すると次のようになる。この図 33 の言込処理は、ホストコンピュータからのファイルデータの言込要求により多数の論理ブロックデータが連続して受信された場合のシーケンシャルライト処理であり、非圧縮キャッシュメモリ 42 からのキャッシュブロックの追出しは、LRU 部 46 に依存せず、新規の言込ブロックを連続的に圧縮回路 48 に供給して圧縮し、圧縮キャッシュメモリ 44 で圧縮データが圧縮グループのサイズである 16 K B となったときに物理デバイスとしてのディスク装置 14 に書き込む場合を例にとっている。

【0144】図 33 において、まずステップ S1 で、ライトキャッシュ制御部 62 はファイルの言込要求に伴ってホストコンピュータ 10 より連続的に送出される論理ブロック番号とブロックデータを受信すると、まずステップ S2 で、非圧縮キャッシュメモリ 42 に記録する。非圧縮キャッシュメモリ 42 に記録された論理ブロックは、直ちに圧縮回路 48 に供給されて圧縮される。

【0145】次に圧縮グループ作成部 52 において 16 K B の圧縮グループにおける位置が判定され、先頭であれば位置フラグ 1 が付加され、途中であれば位置フラグ 2 と途中の順番を示す番号が付加され、更にグループの終端から次のグループの先頭にかかっている位置フラグ 3 が付加される。圧縮グループ作成部 52 で圧縮データに位置フラグが付加されると、ステップ S4 で、図 32 のように論理ブロック番号に対応して、付加した位置フラグをマッピングテーブル 54 に書き込む。続いてステップ S5 で、ホストコンピュータ 10 からのライト要求が終了したか否かチェックし、ライト要求が継続していれば、ステップ S6 で、非圧縮キャッシュメモリ 44 内において 16 K B の圧縮グループの作成完了をチェックする。即ち、圧縮キャッシュメモリ 44 に割り当てた 16 K B のキャッシュブロックに対する圧縮グループ作成部 52 からの位置フラグを付加した圧縮データの格納で、格納データ量が 16 K B に達したか否かチェックしている。

【0146】圧縮グループの格納データ量が 16 K B に達するまでは、ステップ S1 からの処理を繰り返している。圧縮グループのデータ格納量が 16 K B に達すると、ステップ S7 に進み、フリーリンクドリスト 56 からディスク装置 14 の空き状態にある論理セクタを獲得し、圧縮グループをディスク装置に格納する。続いてステップ S8 で、マッピングテーブル 54 にディスク装置 14 に格納した論理グループを構成する論理ブロック番号に対し、獲得した論理セクタ番号を図 32 のように登録する。

【0147】ディスク装置 14 に対する格納が済むと、

(20)

特開平 8-249159

37

ステップS9で、非圧縮キャッシュメモリ42および圧縮キャッシュメモリ44に残っている既に処理の済んだデータを、後続するブロックの処理の空きスペースを確保するために削除する。一方、ステップS5でライト要求の終了を判別した場合には、ステップS10で、圧縮グループの容量が16KBに満たなくとも、フリーリンクドリスト56から論理セクタを獲得してディスク装置14に書き込み、ステップS11で、マッピングテーブル54への獲得した論理セクタ番号の書き込みを行い、ステップS12で、ディスク装置に書き込みが済んだ非圧縮

キャッシュメモリ42および圧縮キャッシュメモリ44の不要なデータを削除して処理を終了する。  
【0148】図34は、図33でのシーケンシャルライト処理によりディスク装置に圧縮して書き込んだファイルデータのシーケンシャルリード処理である。リードキャッシュ制御部64がホストコンピュータ10からのリード要求による論理ブロック番号をステップS1で受信すると、ステップS2でキャッシュヒットの有無を判定する。ステップS2でキャッシュメモリのミスヒットであれば、ステップS3でマッピングテーブル54を参照し、論理ブロック番号から論理セクタ番号を検索し、ス

テップS4で、該当セクタの圧縮データをディスク装置から読み出して圧縮キャッシュメモリ44に転送する。  
【0149】続いてステップS5で、マッピングテーブル54から該当する位置フラグを獲得する。獲得した位置フラグについては、ステップS6で、先頭位置か即ち位置フラグ1か否かチェックされ、ステップS11で、途中か即ち位置フラグ2か否かチェックする。位置フラグから先頭位置であることが判った場合には、圧縮キャッシュメモリ44の圧縮データの先頭から伸長回路50に圧縮データを供給し、次の位置フラグが得られたときに、伸長回路50に対する供給を停止する。

【0150】したがって、圧縮グループの中の位置フラグで指定された先頭部分の圧縮データのみが復元されて、非圧縮キャッシュメモリ42に展開される。ステップS10では、該当ブロックをホストコンピュータに転送する。続いてステップS17で、リード要求の終了をチェックし、リード要求の終了が得られるまで、ステップS1からの処理を繰り返す。

【0151】一方、位置フラグが、ステップS1で、途中を示す位置フラグ2であることが判別された場合には、ステップS12で、要求ブロックが途中位置フラグの何番目かを認識し、該当する位置フラグから圧縮データを取り出して伸長回路50に供給し、次の位置フラグが検出されたときに圧縮データの供給を停止する。これによって、要求ブロックに対応する途中の圧縮データのみが伸長回路50に与えられ、非圧縮キャッシュメモリ42に復元され、ステップS13でホストコンピュータ10に転送することができる。

【0152】次に位置フラグ3が判別された場合には、

38

ステップS14～S16の処理となる。この位置フラグの処理に際し、ステップS4におけるディスク装置からの圧縮グループの読出しにあつては、マッピングテーブルの参照で位置フラグ3が認識され、位置フラグ3の場合には、位置フラグ3を付加した論理セクタと次の論理セクタに亘って圧縮データが格納されていることから、位置フラグ3を認識した場合には、ステップS4において、追続する2つの論理セクタの圧縮データをディスク装置から読み出して圧縮キャッシュメモリ44に転送している。

【0153】このため、ステップS14からの位置フラグ3の処理にあつては、まず位置フラグで特定される圧縮グループの最終位置の圧縮データを伸長回路50に供給して復元し、ステップS5で、次のセクタとなる圧縮データの先頭から次の位置フラグが出るまで伸長して非圧縮キャッシュメモリ42に展開する。最終的に、ステップS16で、非圧縮キャッシュメモリ42に展開されたブロックをホストコンピュータ10に転送する。

【0154】ステップS17でリード要求の終了を判別すると、ステップS18で、圧縮キャッシュメモリ44および非圧縮キャッシュメモリ42に残っているシーケンシャルリードの対象となった不要なデータを削除して、キャッシュの使用領域を確保する。ここで、ステップS2でキャッシュヒットであった場合には、ステップS6で非圧縮キャッシュメモリ42のヒットか否かチェックし、非圧縮キャッシュメモリのヒットであれば、ステップS7で、非圧縮キャッシュメモリ42の要求データをホストコンピュータ10に転送する。また圧縮キャッシュメモリ44のヒットであった場合には、ステップS3、S4によるディスク装置14からの読出しは行わず、直ちにステップS5からの処理を行う。

【0155】尚、圧縮データの先頭に位置フラグを付加して圧縮グループを作成する実施例については、キャッシュメモリにデータを残さないシーケンシャルライトとシーケンシャルリードを例にとっているが、通常のLRI方式に従ったキャッシュメモリを利用した書込処理および読出処理についても、図9の最小物理セクタの整数倍の圧縮ブロックで圧縮グループを形成する場合と全く同様にして適用することができる。

【0156】また上記の実施例にあつては、ディスク装置から圧縮キャッシュメモリ44に読み出した圧縮グループの中の圧縮ブロックの復元について、ホストコンピュータ10の要求ブロックのみを取り出して非圧縮キャッシュメモリ42に展開しているが、圧縮グループ単位に復元して非圧縮キャッシュメモリ42に展開してもよい。これにより非圧縮キャッシュメモリのヒット率を高め、圧縮キャッシュメモリがヒットした場合の伸長処理によるオーバーヘッドを低減することが可能である。

【0157】また上記の実施例は、物理デバイスとしてRAID構成のディスクアレイを例にとっているが、ア

(21)

特開平8-249159

39

レイ構成をもたない通常のディスク装置であってもよい。また、物理デバイスは磁気ディスク装置以外に、光ディスク装置、半導体メモリ装置などの通常の外部記憶装置を用いることができる。

#### 9. 他の動作環境

図35は本発明のディスク制御装置の他の動作環境である。図2の実施例では上位のホストコンピュータに接続された場合を例にとっているが、図35の動作環境にあっては、本発明のディスク制御装置をディスクサーバーとして使用したことを特徴とする。

【0158】図35において、ディスクアレイを構成するディスク装置14-1～14-6を備えた本発明のディスク制御装置12は、ローカルエリアネットワーク（以下「LAN」という）200を介してホストコンピュータ10に接続される。LAN200には、ホストコンピュータ10以外にワークステーション202、パソコン204、更にプリンタ206が接続される。

【0159】このため本発明のディスク制御装置12は、LAN200に接続したクライアントであるホストコンピュータ10、ワークステーション202、及びパソコン204からの要求を受け付けて処理するディスクサーバーとして機能する。更に、ホストコンピュータ10は通信制御装置208を介して外部のネットワーク通信回線210と接続されており、このネットワーク接続によりLAN200を越えた多くのクライアントからの要求を受付けるディスクサーバーとして機能できる。即ち、ネットワーク接続により多数のパソコン、ホストと接続可能となるばかりでなく、地球規模で接続された場合、昼夜の別なくアクセスされ、データの記録、再生が続けられることになる。

【0160】従って、このような接続形態のもとでは、常にデータをディスク装置に効率よく格納することが必死であり、途中でディスク装置のセグメントクリーニングによる装置の停止は許されない。またスタントアロン型のディスク制御装置であっても、ユーザが装置の使用中にセグメントクリーニングが入ることになり、ユーザ作業が中断されることになる。同様に、大型のホストコンピュータに接続されたディスク制御装置の場合にも、セグメントクリーニング中は作業が中断されることになる。

【0161】いずれの場合も、セグメントクリーニングは装置の動作中には、処理が中断されることになり、装置の稼働効率は向上しない。このような場合、本発明のディスク制御装置は、処理の中断を伴うセグメントクリーニング処理を無くすものであり、特に昼夜の別なくアクセスされる地球規模のネットワークと接続される無停止型ディスクサーバーとしての効果は極めて大きい。

#### 【0162】

【発明の効果】以上説明してきたように本発明によれば、外部記憶システムの制御装置でデータ圧縮を行って

40

いることから、上位装置のデータ圧縮に関する負担をなくし、データ圧縮機能を設けることで既存の外部記憶システムの容量を引き上げて、低コストのシステムで実現できる。

【0163】また、記憶システムの制御装置で上位装置からの固定長データを圧縮した場合、圧縮データ長がデータの性質により様々変化するが、複数の圧縮データをまとめて1つの固定長の圧縮グループを作り、圧縮グループ単位にディスク装置に対し記録再生することで、可変長データとなる圧縮データの固定長記録を可能とし、ディスク装置の内部的なフラグメントを最小限に抑えることができる。このため、可変長データの格納で従来生じていたディスク装置のセグメントクリーニングが不要となり、セグメントクリーニングによるデバイスビジーが発生しないことでアクセス性能を向上させることができる。

#### 【図面の簡単な説明】

【図1】本発明の原理説明図

【図2】本発明が適用される動作環境のブロック図

【図3】図2のシステムを実現するハードウェア構成のブロック図

【図4】図3の圧縮モジュールの詳細ブロック図

【図5】本発明の機能ブロック図

【図6】キャッシュメモリの説明図

【図7】キャッシュテーブルの説明図

【図8】非圧縮キャッシュメモリと圧縮キャッシュメモリの説明図

【図9】4KBの圧縮データの説明図

【図10】圧縮グループの候補リストの説明図

【図11】マッピングテーブルの説明図

【図12】フリーリンクドリストの説明図

【図13】圧縮からディスク格納までの流れを示した説明図

【図14】ミスヒット時の書込処理の説明図

【図15】非圧縮キャッシュにヒットした時の書込処理の説明図

【図16】圧縮キャッシュにヒットした時の書込処理の説明図

【図17】キャッシュ追出し時のライトバック処理の説明図

【図18】ライトキャッシュ制御部のライト処理のフローチャート

【図19】図18のマッピングテーブル更新処理のフローチャート

【図20】キャッシュにミスヒット時の読出処理の説明図

【図21】リードキャッシュ制御部のリード処理のフローチャート

【図22】二重化構成のブロック図

【図23】二重化構成のライト処理のフローチャート

(22)

特開平8-249159

41

42

【図24】図23の更新処理1, 2のフローチャート  
 【図25】図24のマッピングテーブル更新処理のフローチャート  
 【図26】図23の更新処理3, 4のフローチャート  
 【図27】図26の二重化処理のフローチャート  
 【図28】図23の更新処理5のフローチャート  
 【図29】二重化構成のリード処理のフローチャート  
 【図30】ダミーデータの付加を必要としない圧縮グループ形成の説明図  
 【図31】図30の圧縮グループを取出した説明図 10  
 【図32】圧縮グループのLRI追出し時に作成されるマッピングテーブルの説明図  
 【図33】位置フラグを付加するライト処理のフローチャート  
 【図34】位置フラグを利用したリード処理のフローチャート  
 【図35】本発明の他の動作環境のブロック図  
 【符号の説明】  
 10: 上位装置 (ホストコンピュータ)  
 12, 12-1, 12-2: 磁気ディスク制御装置 20  
 14: 物理デバイス  
 14-1~14-6: ディスク装置  
 16, 16-1, 16-2: MPU  
 18, 18-1, 18-2: ROM  
 20, 20-1, 20-2: RAM  
 22, 22-1, 22-2: ホストインタフェース制御装置  
 24, 24-1, 24-2: ディスクインタフェース制御装置  
 26, 26-1, 26-2: キャッシュ制御部 30  
 28, 28-1, 28-2: キャッシュメモリ  
 30, 30-1, 30-2: ヒット判定回路  
 32, 32-1, 32-2: 圧縮伸長回路  
 34, 34-1, 34-2: 内部バス  
 36, 36-1, 36-2: チャンネルバス  
 38, 38-1, 38-2: デバイスバス  
 40: キャッシュテーブル  
 42: 非圧縮キャッシュメモリ (非圧縮キャッシュ領域)  
 44: 圧縮キャッシュメモリ (圧縮キャッシュ領域) 40  
 46, 55: LRI部  
 48: 圧縮回路  
 50: 伸長回路  
 52: 圧縮グループ作成部  
 54: マッピングテーブル

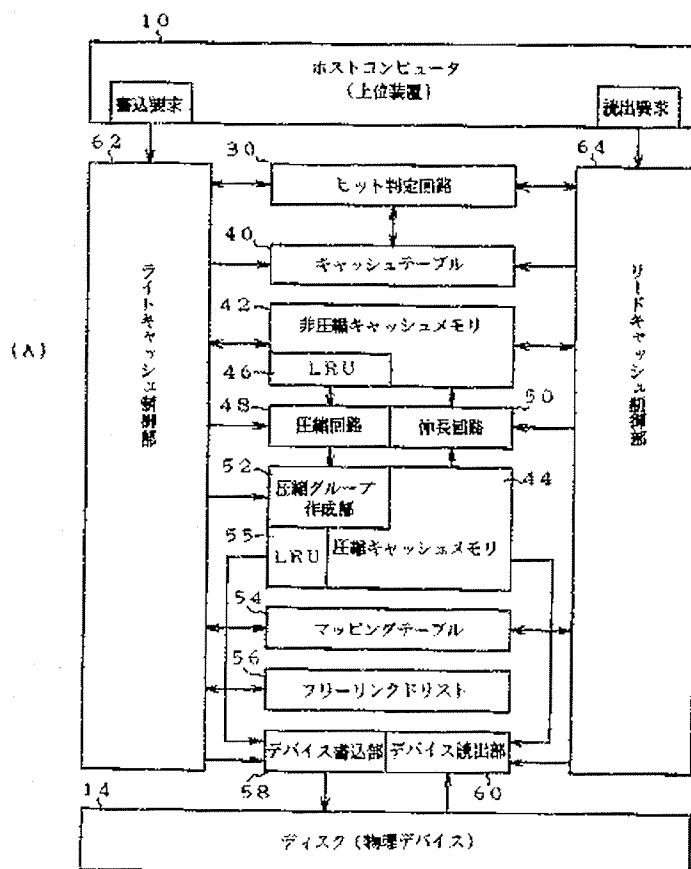
56: フリーリンクドリスト  
 58: デバイス書込部  
 60: デバイス読出部  
 62: ライトキャッシュ制御部  
 64: リードキャッシュ制御部  
 68-1~68-n: キャッシュブロック (キャッシュセグメント)  
 70-1~70-n: 圧縮グループ (圧縮クラスタ)  
 72-1~72-4: 圧縮データ  
 74: サイズリスト  
 76-1~76-7: 組合せ候補  
 78: セクタ情報  
 80: 論理セクタ  
 90: 位置フラグ  
 100: ホストIF用プロセッサモジュール  
 102: ホストIFモジュール  
 104: データ圧縮伸長モジュール  
 106: デバイスIF用プロセッサモジュール  
 108-1~108-6: デバイスIFモジュール  
 110, 136: メインプロセッサ  
 112, 138: コプロセッサ  
 114, 140: フラッシュメモリ  
 116, 142: プロセッササポートロジック  
 118, 144: システムストレージ  
 122: ホストバッファIF  
 124: キャッシュメモリ  
 125: バックアップ電源  
 126: 上位インタフェース部  
 128: 伸長用下位インタフェース部  
 130: 伸長データ格納用バッファメモリ  
 132: 圧縮データ格納用バッファメモリ  
 134-1~134-4: 圧縮伸長ロジック  
 146: デバイスバッファIF  
 148: バッファメモリ  
 150: デバイスIF  
 152: ブリッジバッファIF  
 154: 周辺ロジック  
 156: プロセッサバスブリッジ  
 200: ローカルエリアネットワーク  
 202: ワークステーション  
 204: パソコン  
 206: プリンタ  
 208: 通信制御装置  
 210: 通信回線

(23)

特開平 8-249159

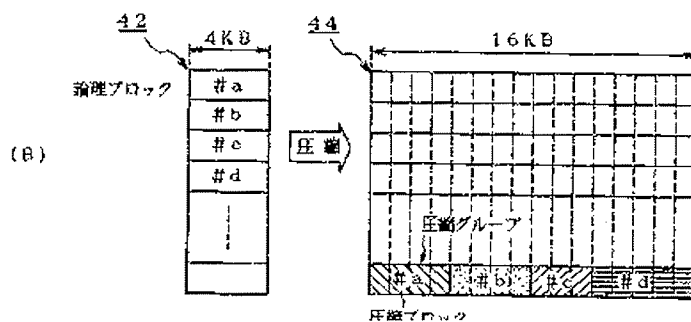
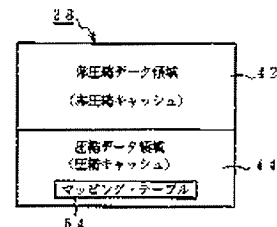
【図1】

本発明の原通説明図



【図6】

キャッシュメモリの説明図

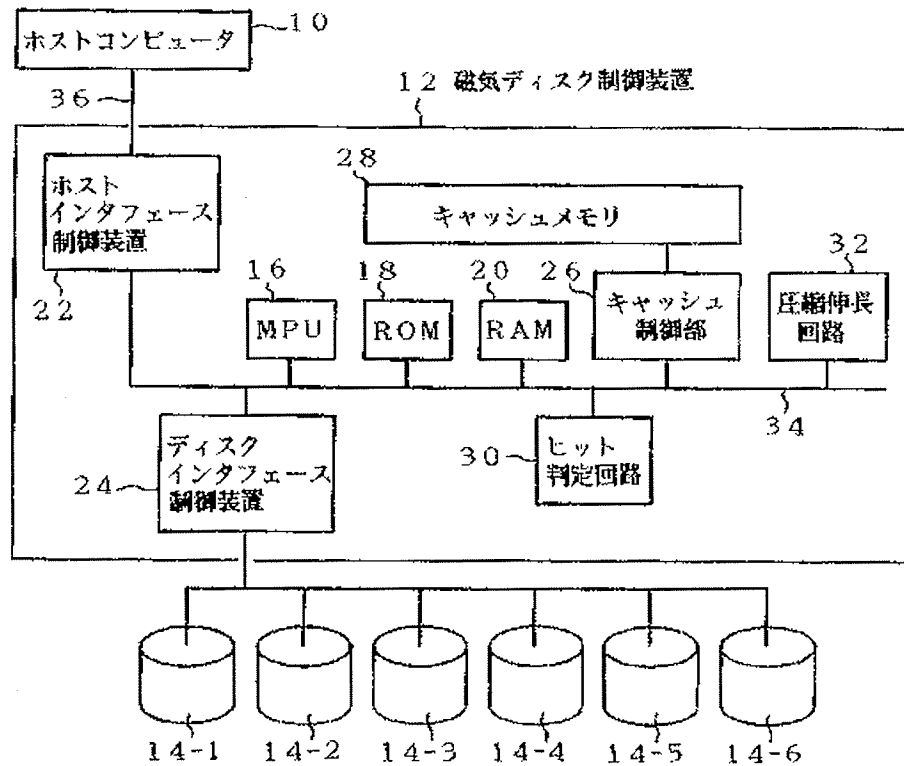


(24)

特開平8-249159

【図2】

本発明が適用される動作環境のブロック図



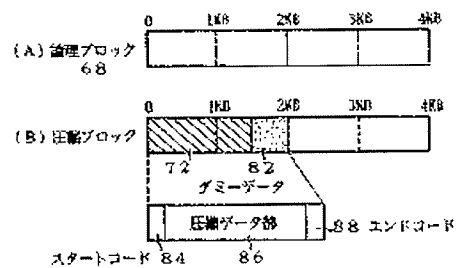
【図7】

キャッシュテーブルの説明図

	40-1	40-2	40-3	40-4	40-5
	論理ブロック 番号(アドレス)	キャッシュ内 のアドレス	データ長	圧縮フラグ	更新フラグ
キャッシュ テーブル 40	・	・	・	・	・
	・	・	・	・	・
	・	・	・	・	・
	・	・	・	・	・
	・	・	・	・	・
	・	・	・	・	・
	・	・	・	・	・
	・	・	・	・	・
	・	・	・	・	・

【図9】

4KBの圧縮データの説明図

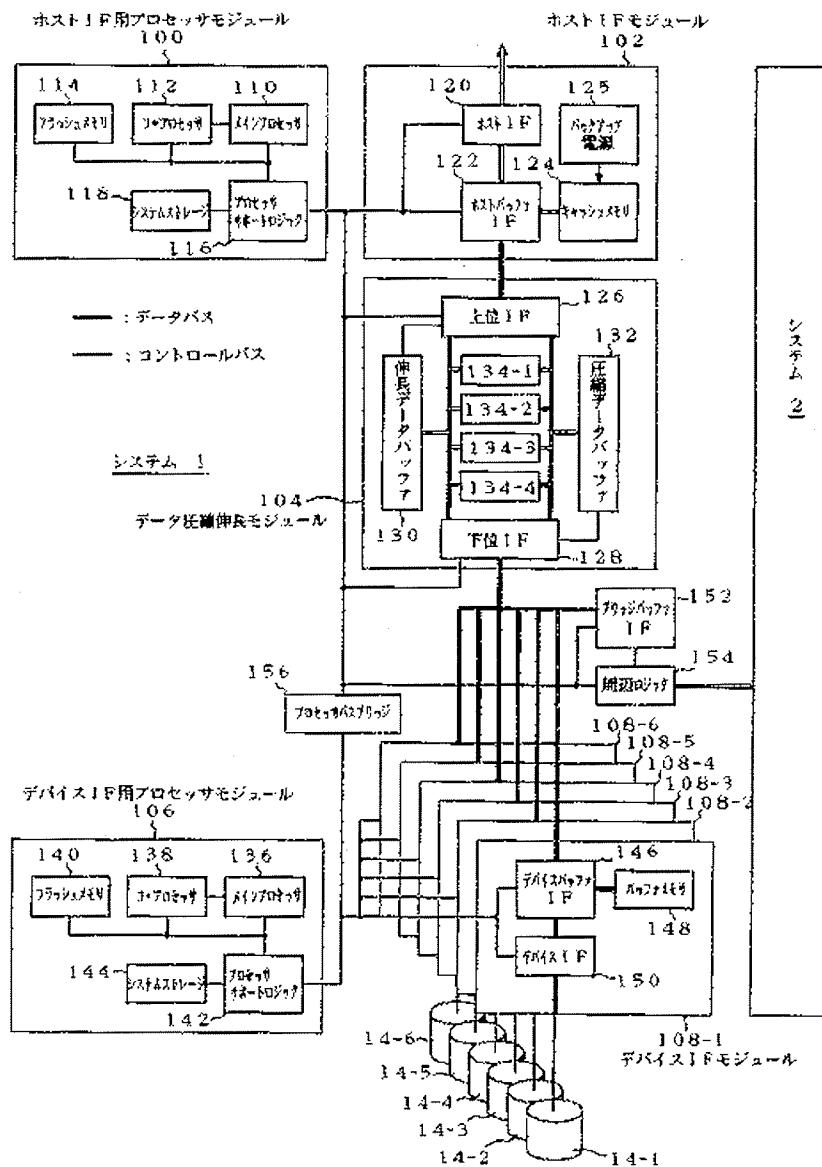


(25)

特開平8-249159

【図3】

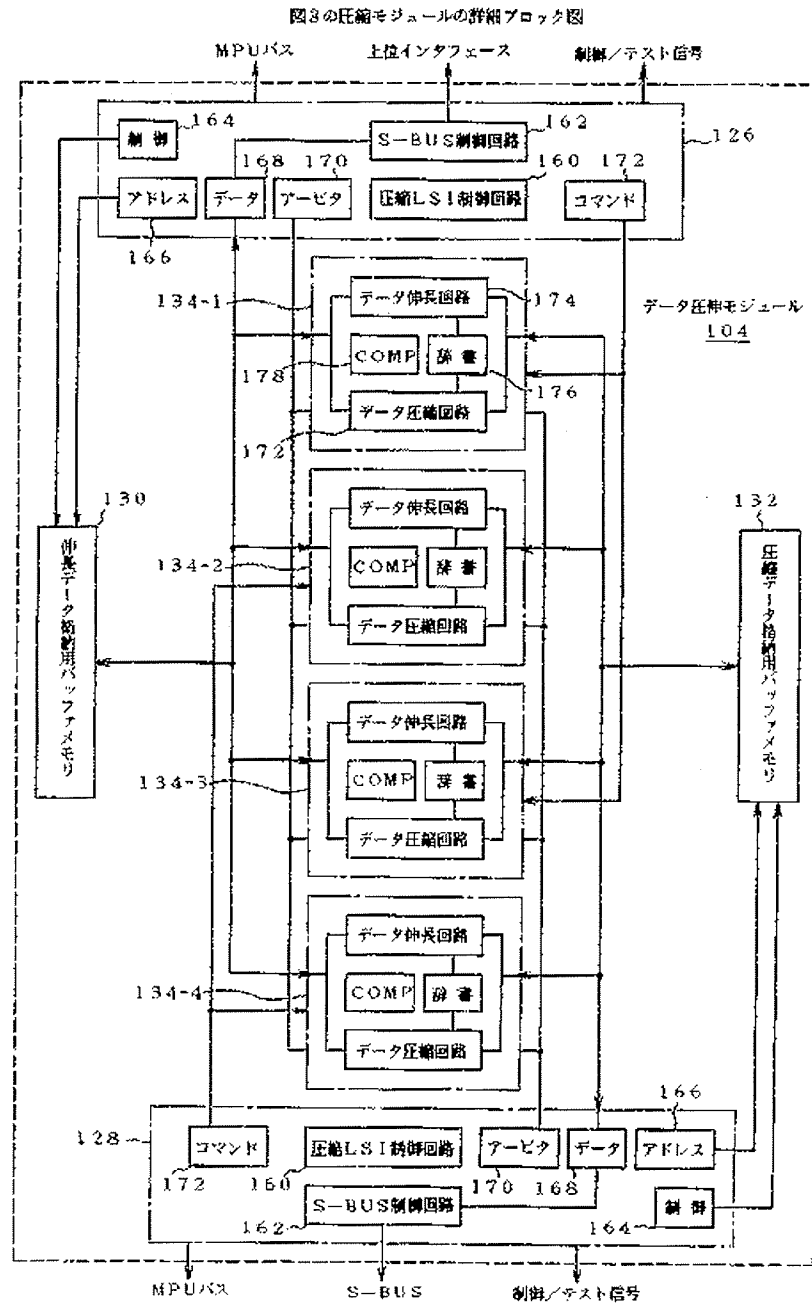
図2のシステムを実現するハードウェア構成のブロック図



(26)

特開平8-249159

【図4】

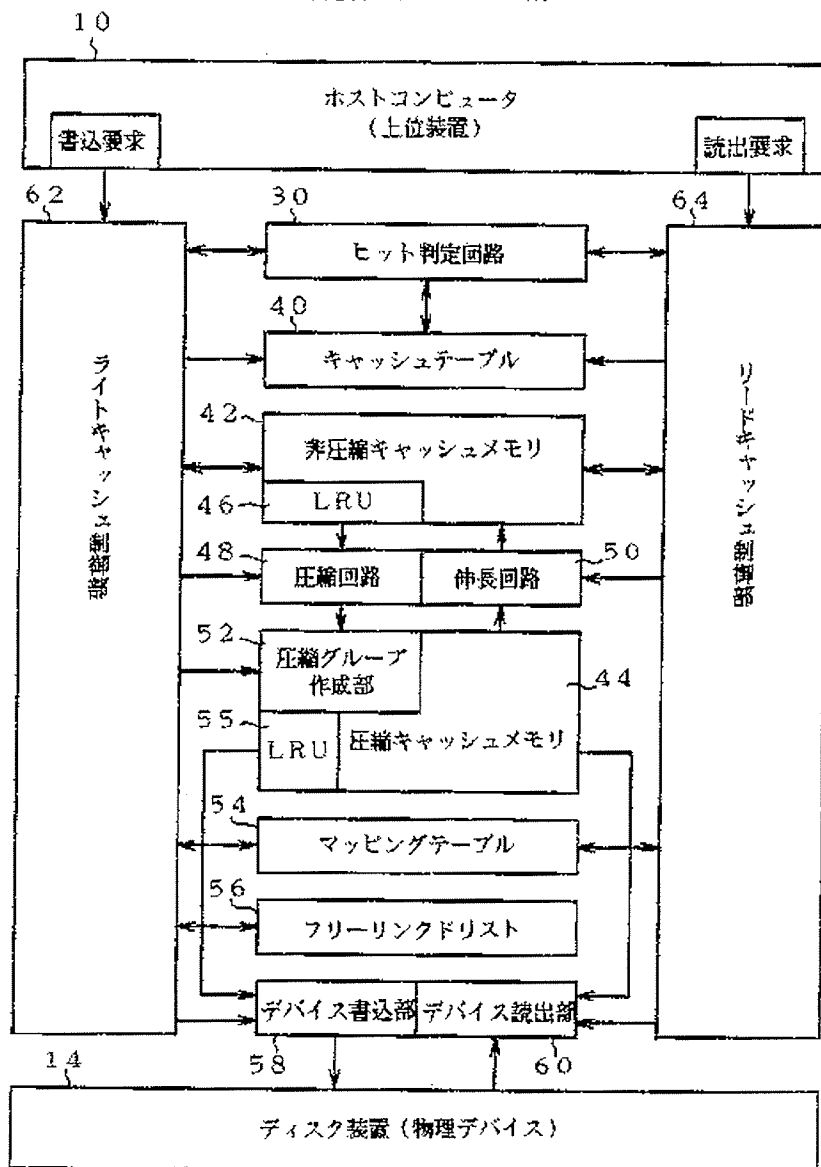


(27)

特開平8-249159

【図5】

本発明の機能ブロック図

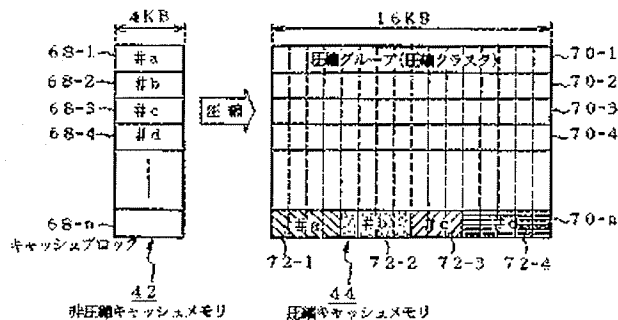


(28)

特開平 8-249159

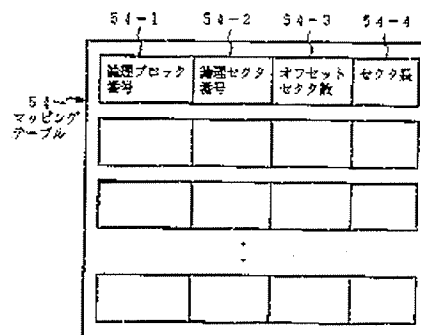
【図 8】

非圧縮キャッシュメモリと圧縮キャッシュメモリの説明図



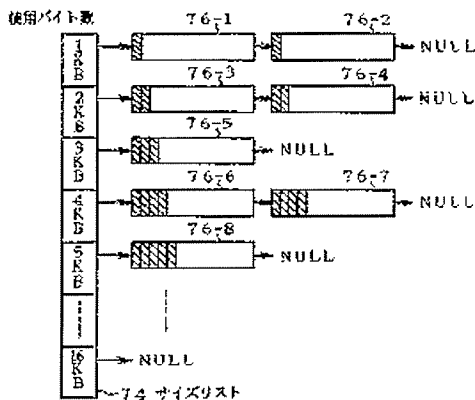
【図 11】

マッピングテーブルの説明図



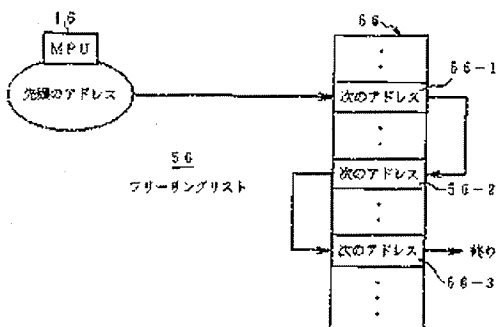
【図 10】

圧縮グループの格納リストの説明図



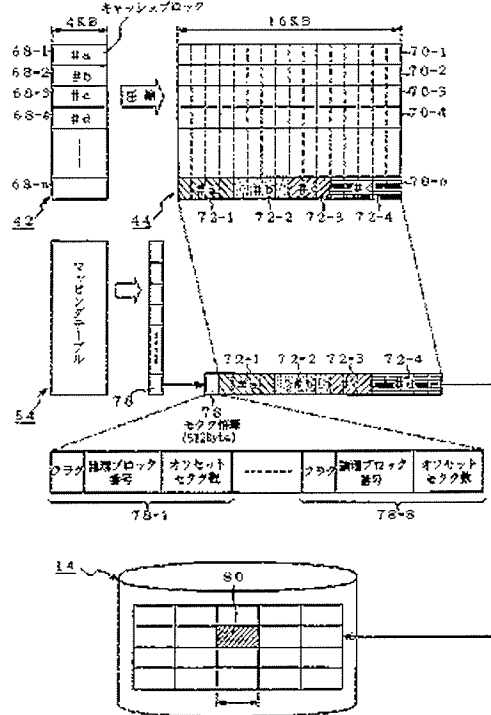
【図 12】

フリーリンクドリストの説明図



【図 13】

圧縮からディスク格納までの流れを示した説明図

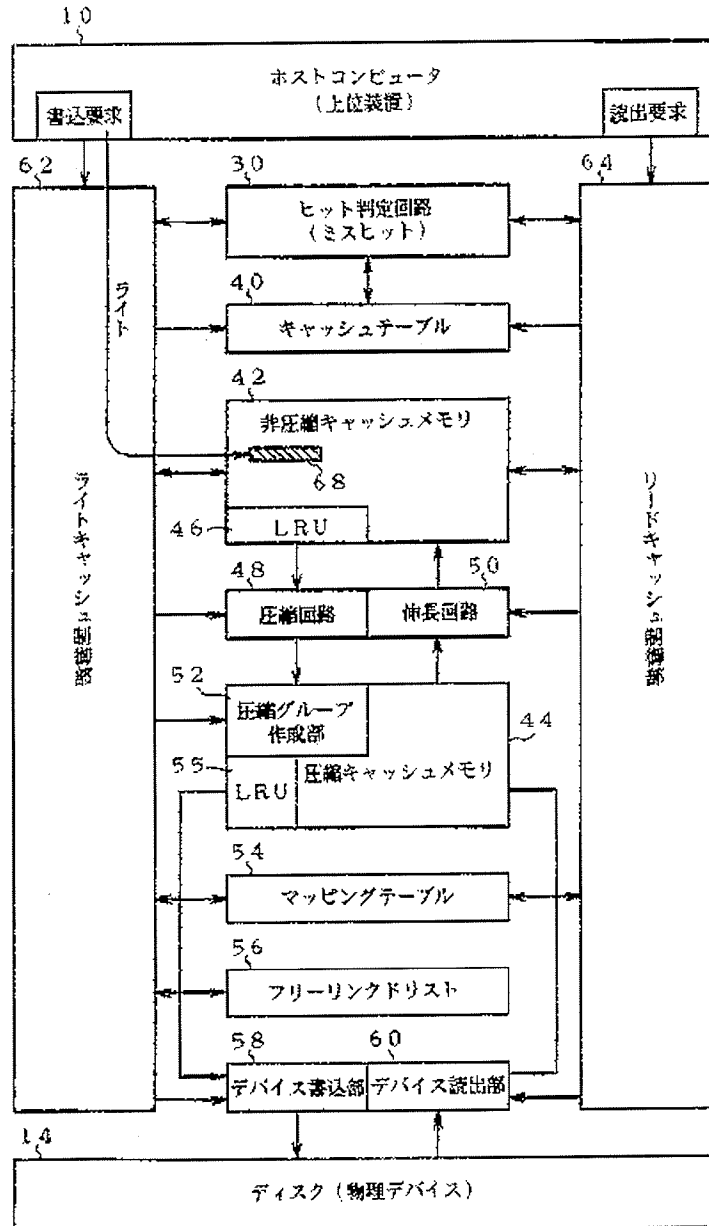


(29)

特開平8-249159

【図14】

ミスヒット時の書込処理の説明図

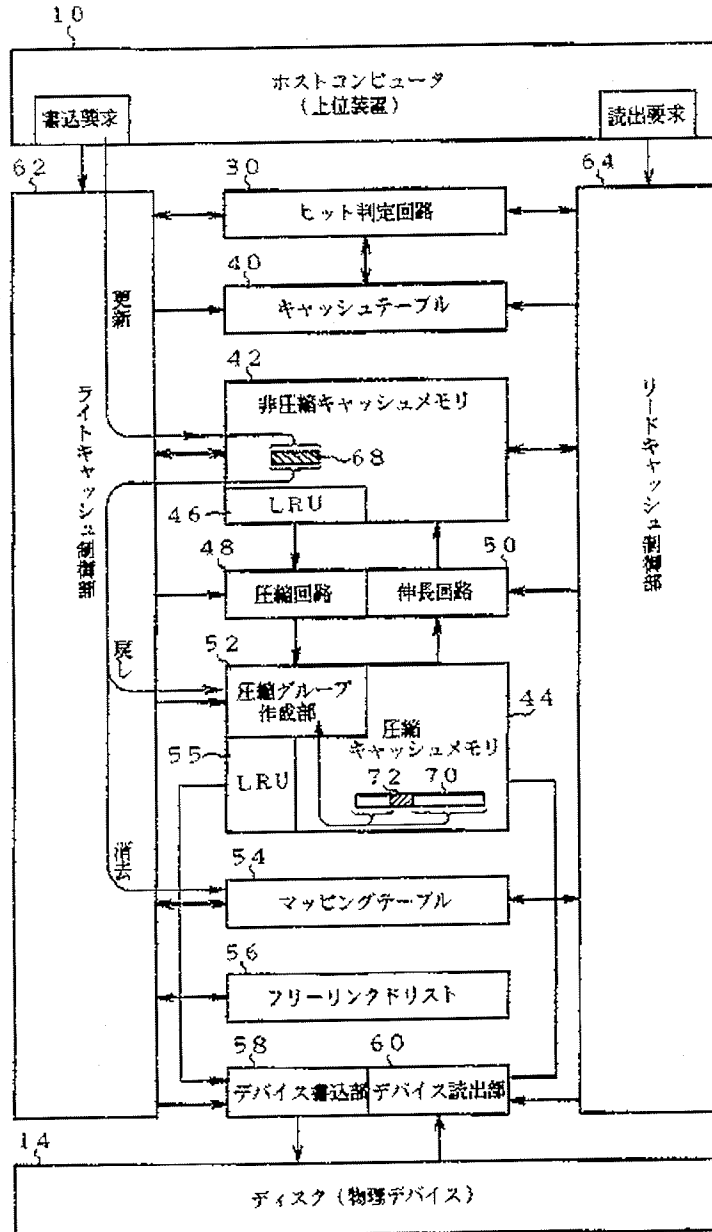


(30)

特開平8-249159

【図15】

非圧縮キャッシュにヒットした時の書込処理の説明図

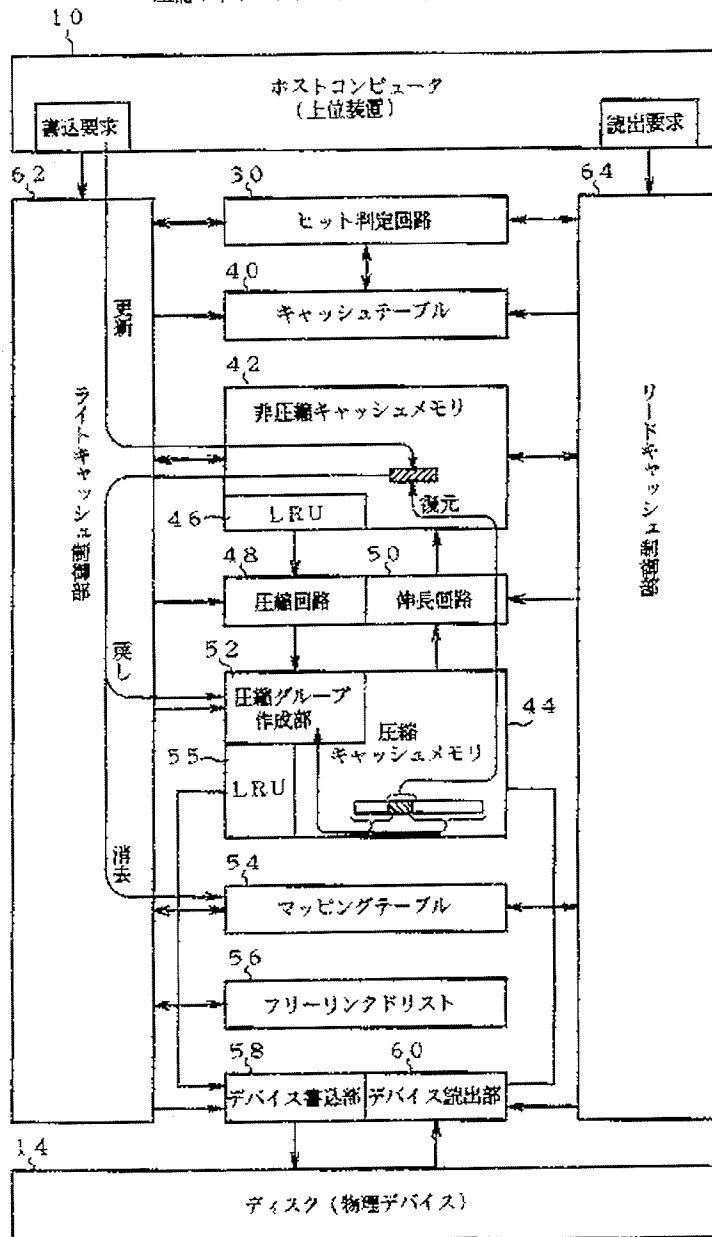


(31)

特開平8-249159

【図16】

圧縮キャッシュにヒットした時の書込処理の説明図

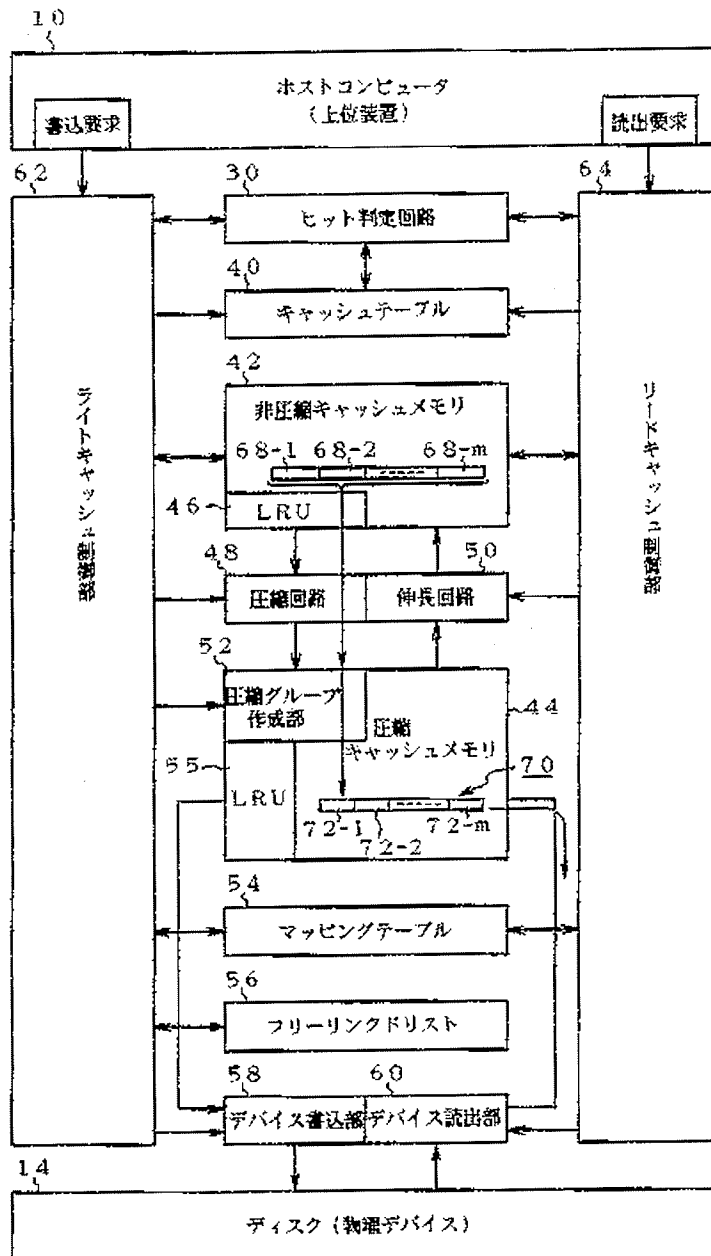


(32)

特開平8-249159

【図17】

キャッシュ追出し時のライトバック処理の説明図

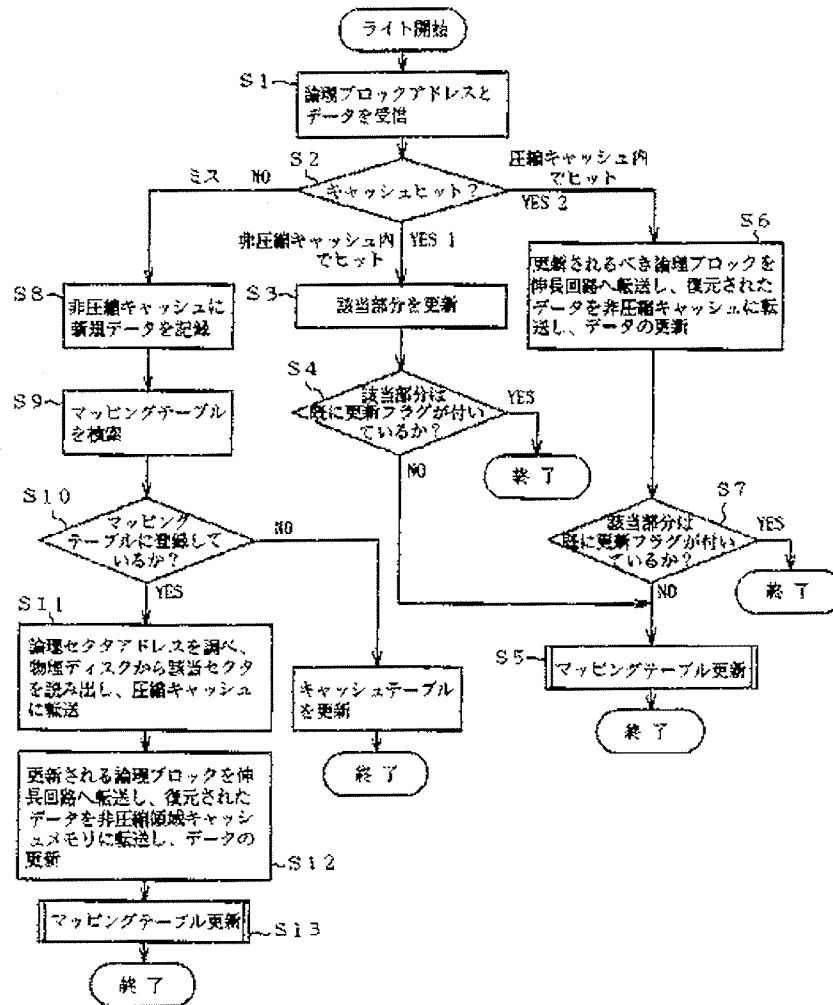


(33)

特開平 8-249159

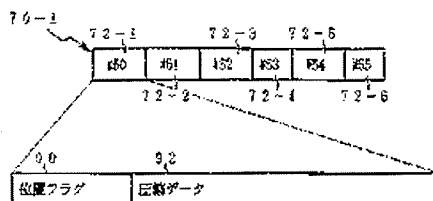
【図18】

ライトキャッシュ制御部のライト処理のフローチャート



【図31】

図30の圧縮グループを取出した説明図



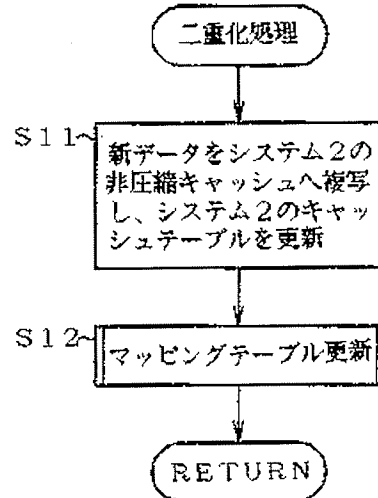
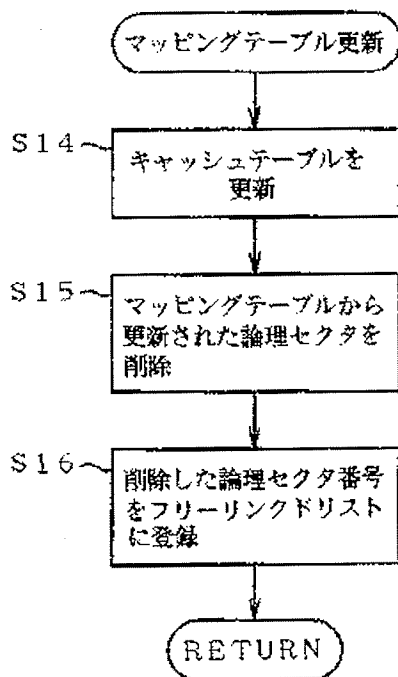
(34)

特開平 8-249159

【図19】

【図27】

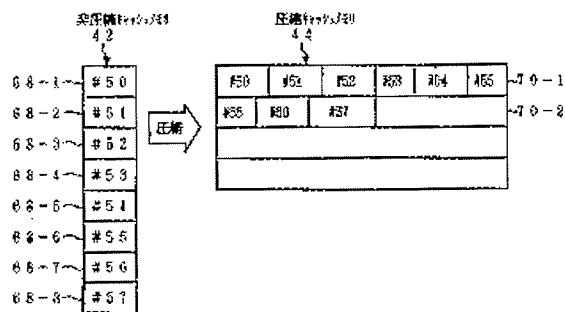
図18のマッピングテーブル更新処理のフローチャート 図26の二重化処理のフローチャート



【図30】

【図32】

ゲームデータの付加を必要としない圧縮グループ形成の説明図



圧縮グループのしり出し処理時に作成されるマッピングテーブルの説明図

64-1	54-2	64-5
論理ブロック番号	論理セクタ番号	位置フラグ
#50	10	1
#51	10	2-1
#52	10	2-2
#53	10	2-3
#54	10	2-4
#55	10	3
#56	11	2-5
#57	11	2-6

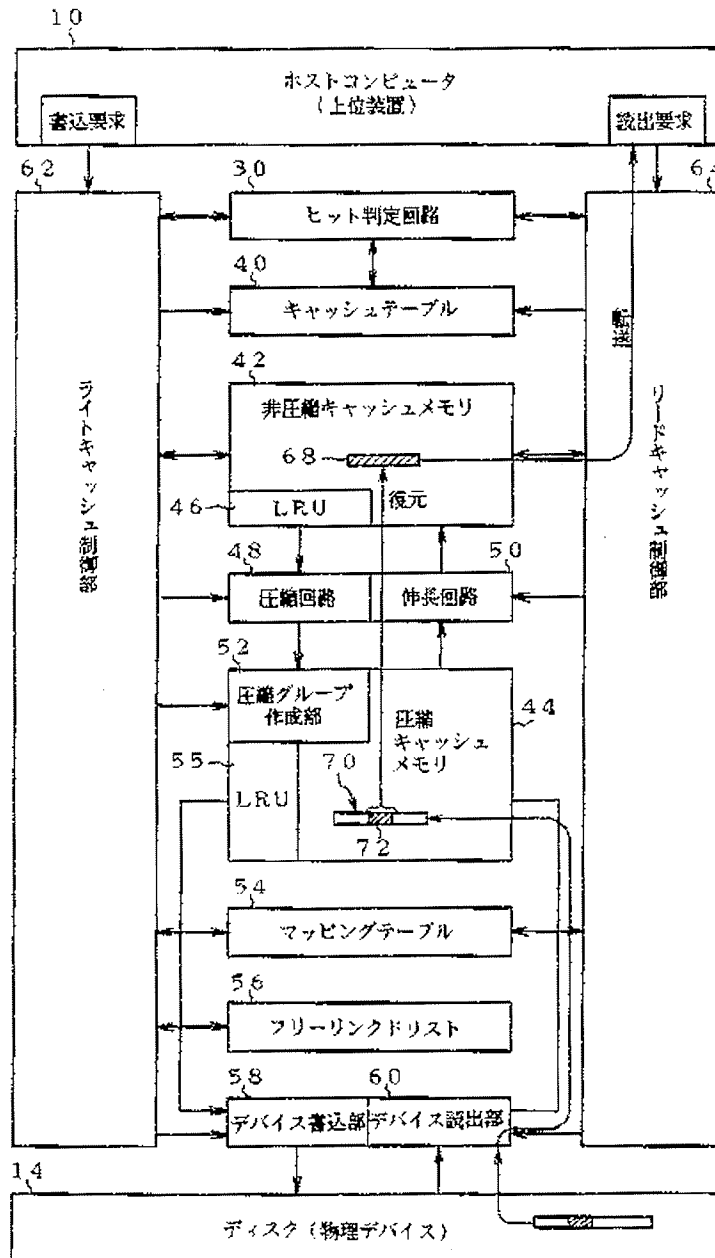
マッピングテーブル

(35)

特開平8-249159

【図20】

キャッシュにミスヒット時の読出処理の説明図

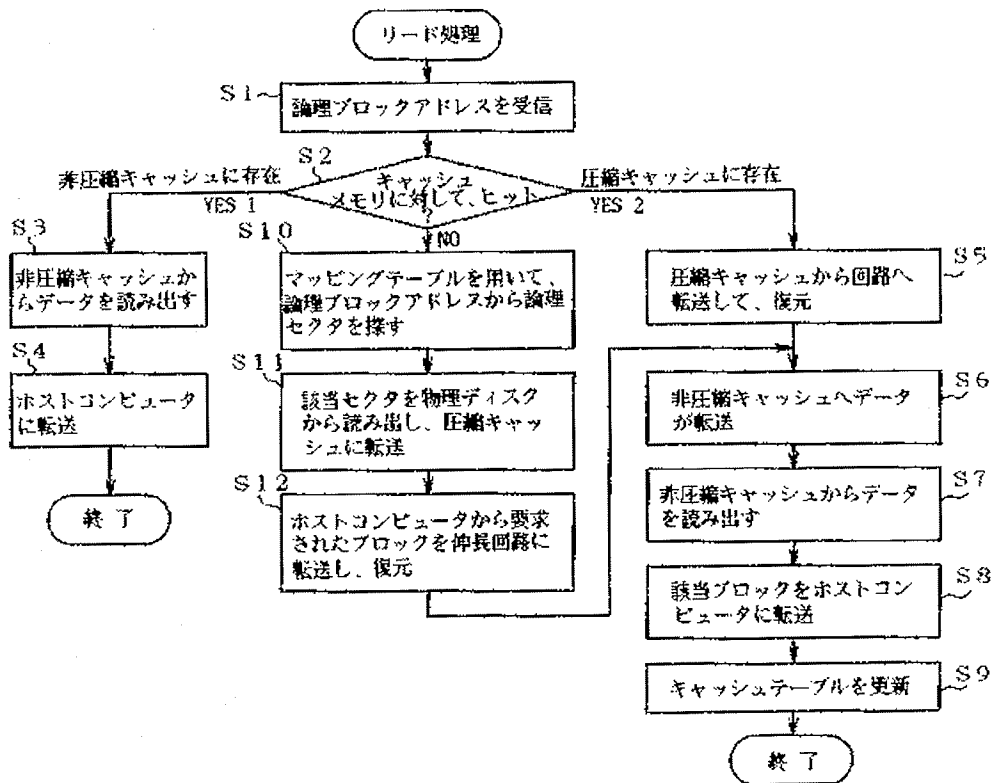


(36)

特開平8-249159

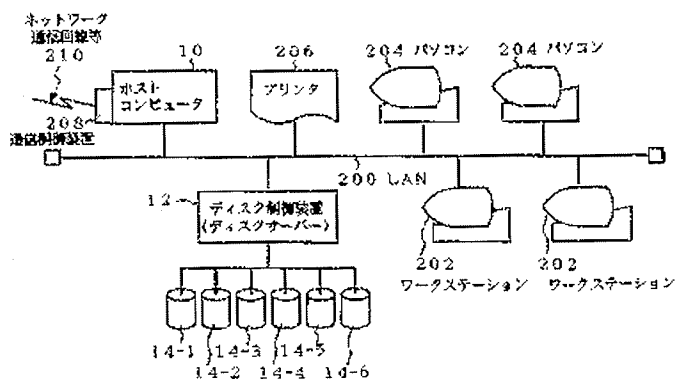
【図21】

リードキャッシュ制御部のリード処理のフローチャート



【図35】

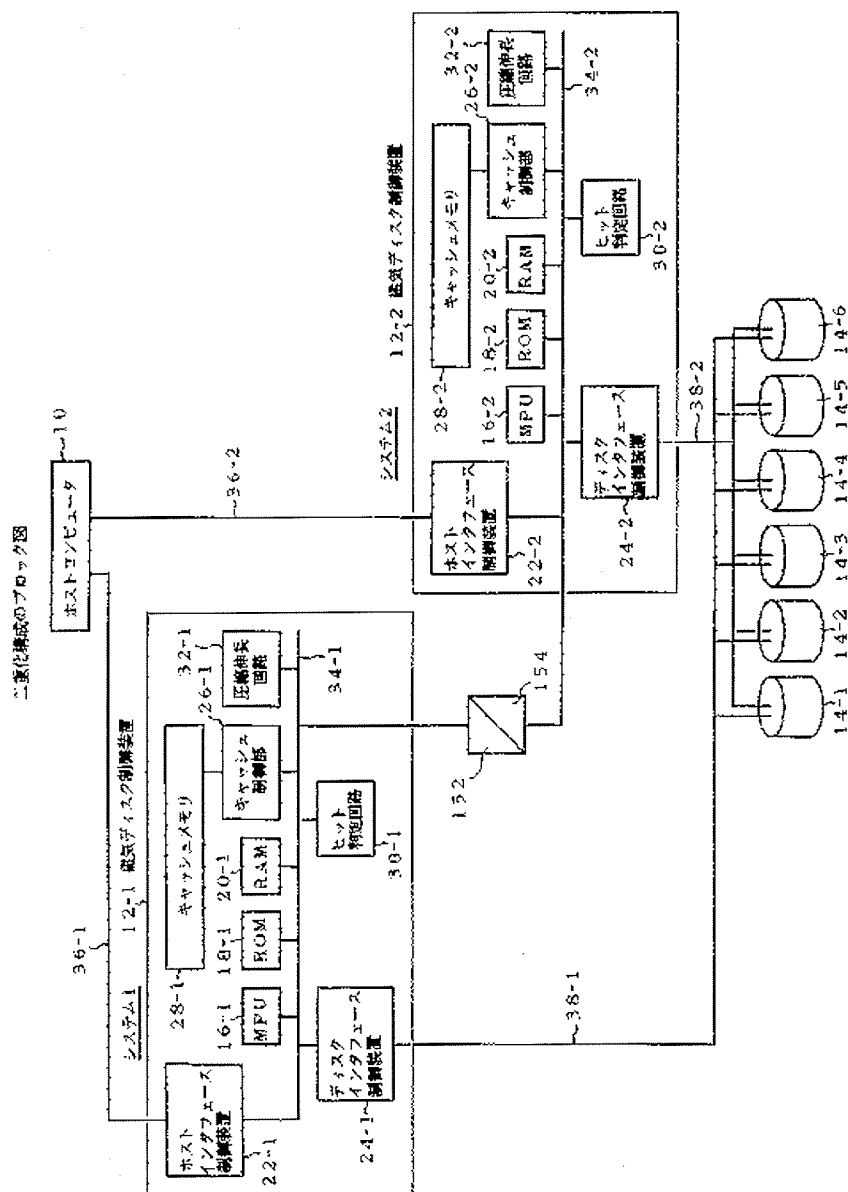
本発明の他の動作態様のブロック図



(37)

特開平8-249159

【図22】

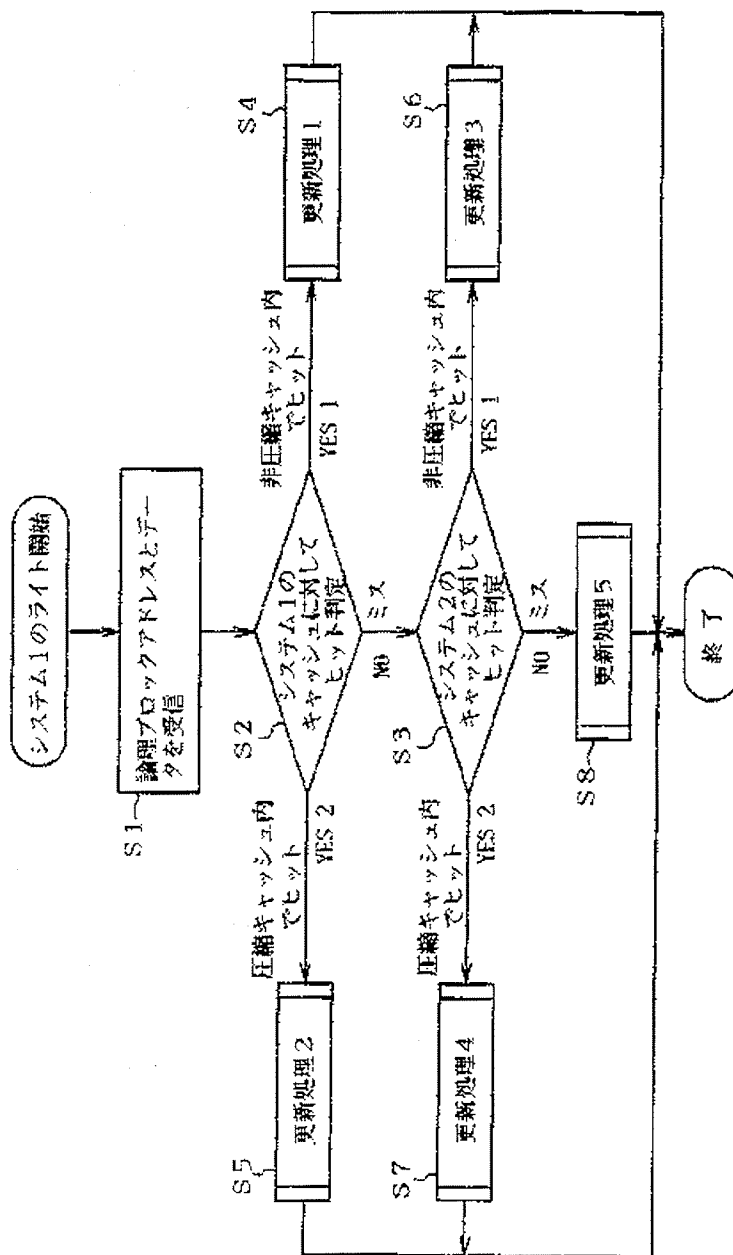


(38)

特開平8-249159

【図23】

## 二重化構成のライト処理のフローチャート

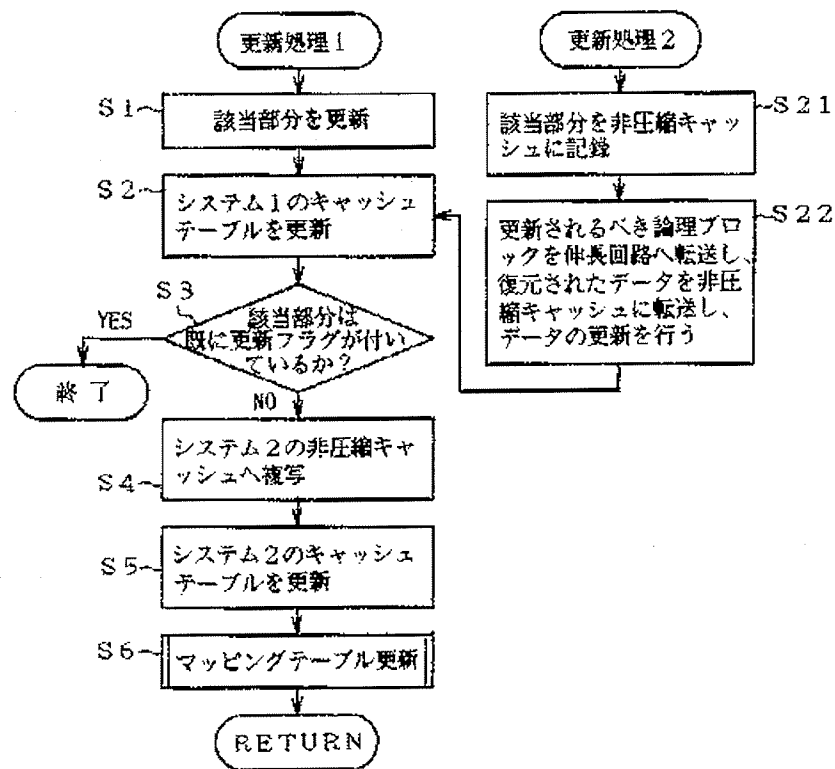


(39)

特開平 8-249159

【図24】

図23の更新処理1, 2のフローチャート

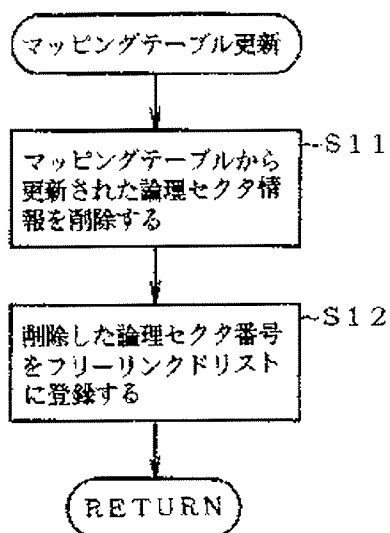


(40)

特開平 8-249159

【図 25】

図 24 のマッピングテーブル更新処理のフローチャート

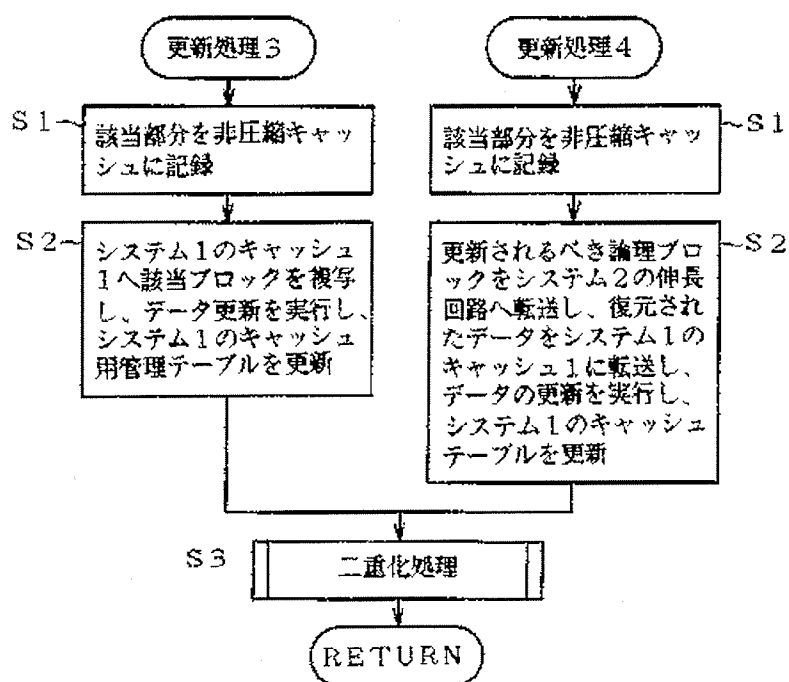


(41)

特開平8-249159

【図26】

図23の更新処理3、4のフローチャート

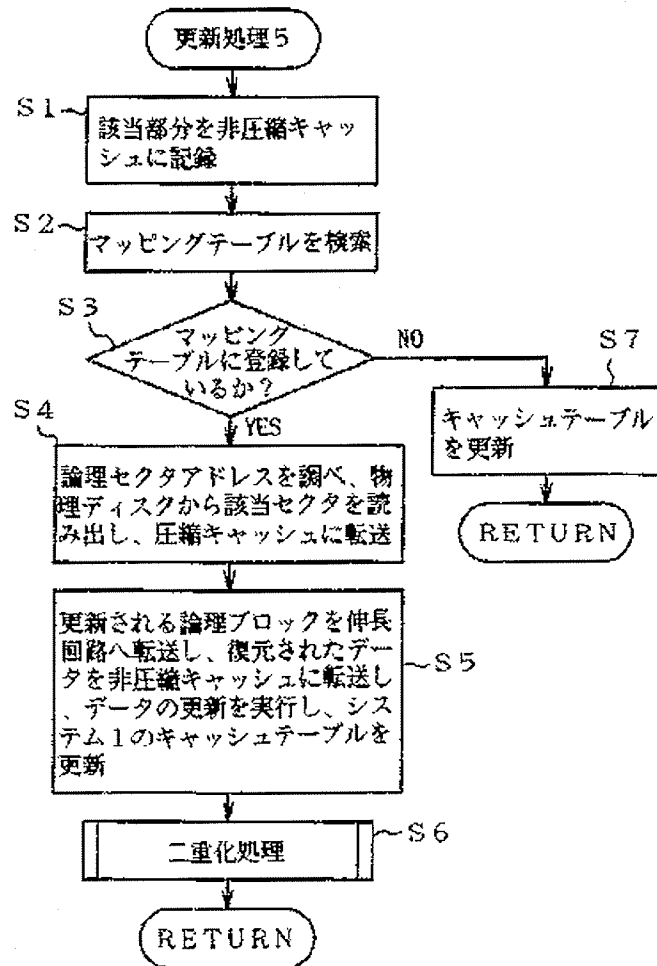


(42)

特開平8-249159

【図28】

図23の更新処理5のフローチャート

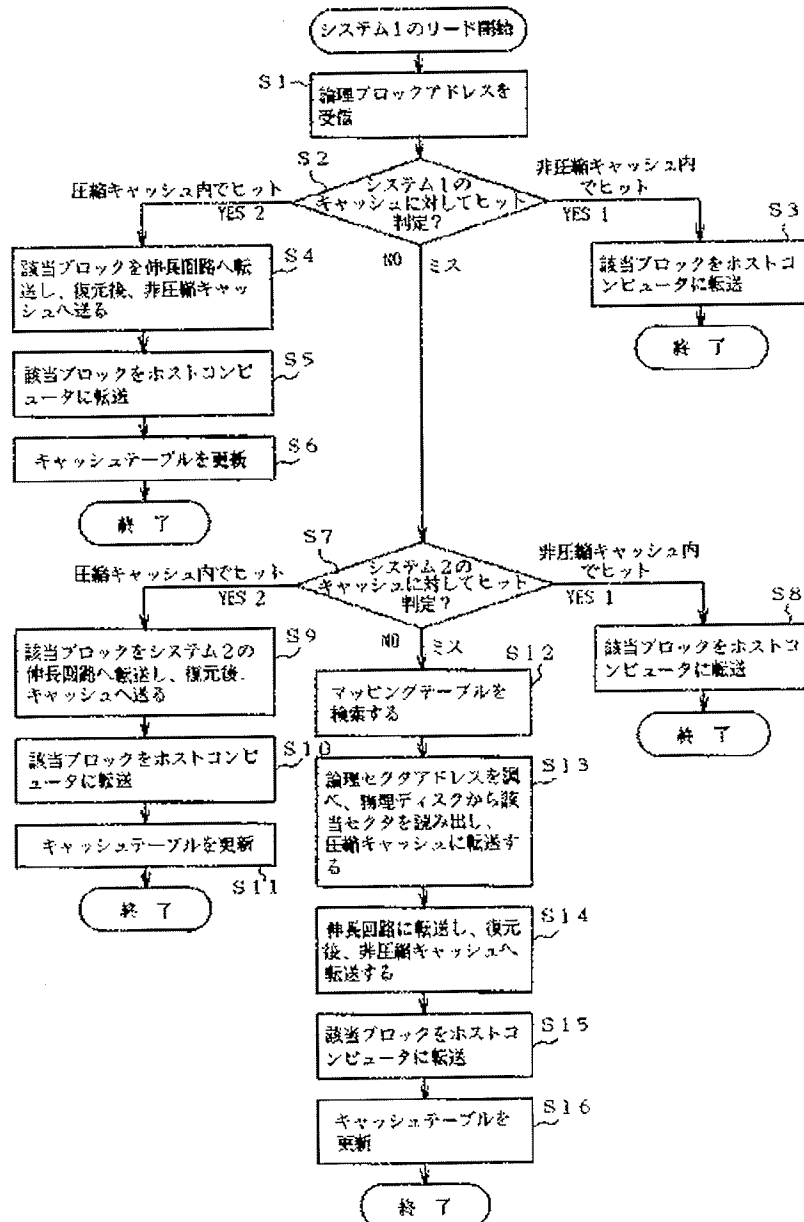


(43)

特開平8-249159

【図29】

二重化構成のリード処理のフローチャート

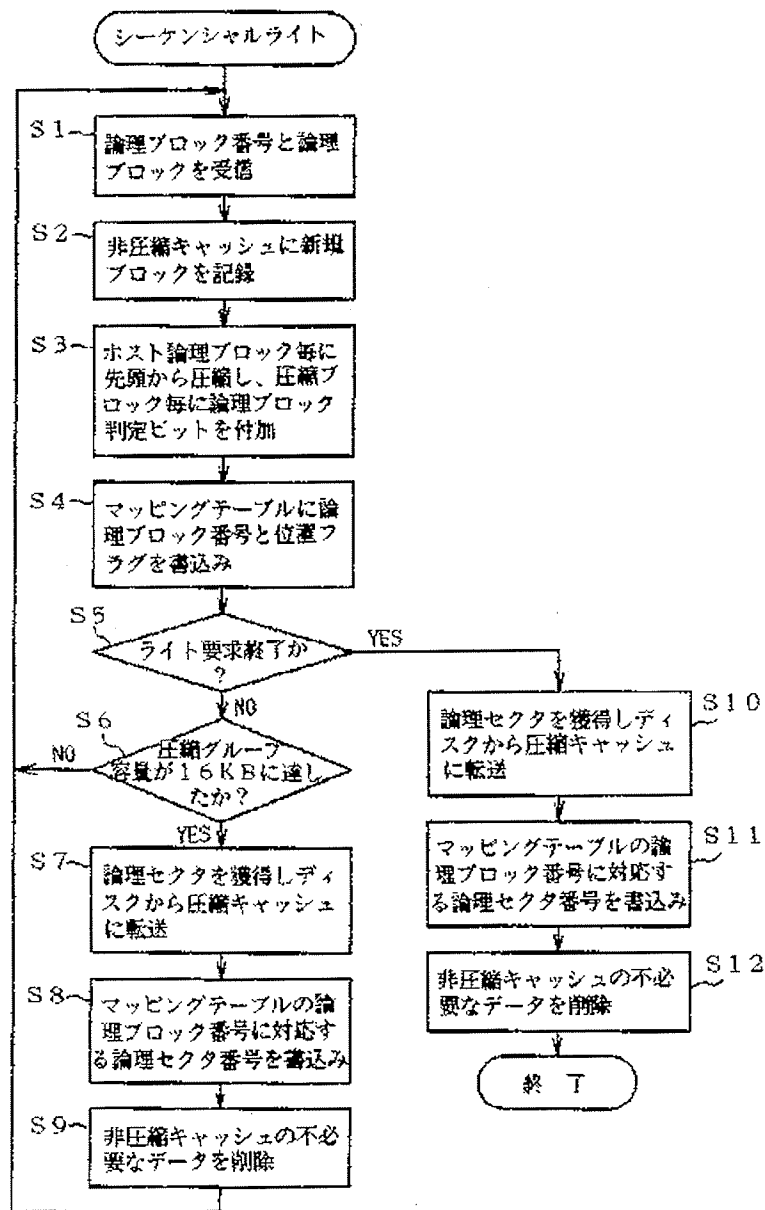


(44)

特開平 8-249159

【図33】

位置フラグを付加するライト処理のフローチャート

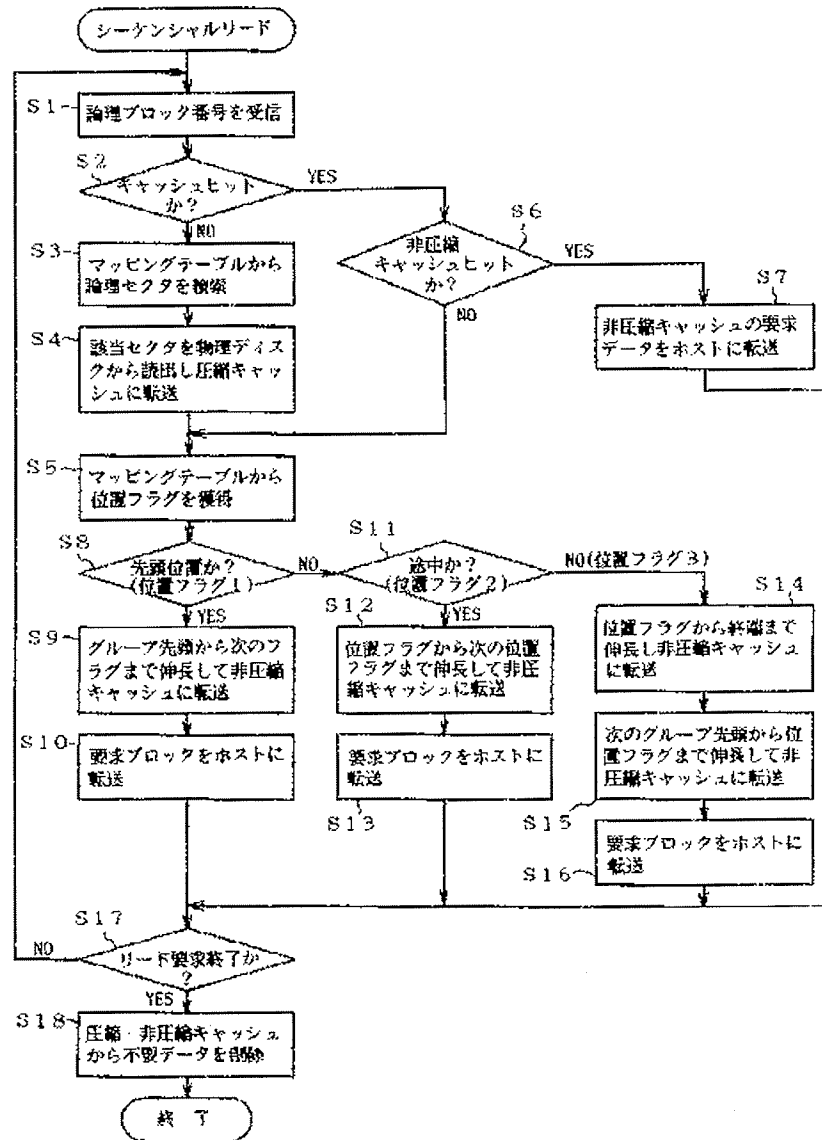


(45)

特開平8-249159

【図34】

位置フラグを利用したリード処理のフローチャート



フロントページの続き

(72)発明者 金子 信  
 神奈川県川崎市中原区上小田中1015番地  
 富士通株式会社内

(72)発明者 本村 茂  
 兵庫県加東郡社町佐保35番(番地なし)  
 富士通周辺機株式会社内

(46)

特開平8-249159

(72)発明者 前田 英友  
兵庫県加東郡社町佐保35番(番地なし)  
富士通周辺機株式会社内